

A

jic957 U.S. PTO

A standard linear barcode consisting of vertical black bars of varying widths on a white background.

Docket Number	4-31192/CIP
---------------	-------------

**FILING BY "EXPRESS MAIL" UNDER 37 CFR 1.10**

EL489123769US  
Express Mail Label Number

November 17, 2000  
Date of Deposit

Address to: Commissioner for Patents  
Box Patent Application  
Washington, DC 20231

## UTILITY PATENT APPLICATION TRANSMITTAL AND FEE SHEET

Transmitted herewith for filing under 37 CFR §1.53(b) is the utility patent application of

Applicant (or identifier): PIPPIG ET AL.

Title: SELECTABLE CELL SURFACE MARKER GENES

**Enclosed are:**

1. ☒ Specification (Including Claims and Abstract) - 55 pages
2. ☒ Drawings - 6 figures; 8 sheets
3. ☒ Unexecuted Declaration and Power of Attorney (original or copy)
4. ☐ Microfiche Computer Program (appendix)
5. ☐ Nucleotide and/or Amino Acid Sequence Submission
  - ☒ Computer Readable Copy
  - ☒ Paper Copy
  - ☒ Statement Verifying Identity of Above Copies
6. ☐ Preliminary Amendment
7. ☐ Assignment Papers (Cover Sheet & Document(s))
8. ☐ English Translation of
9. ☐ Information Disclosure Statement
10. ☐ Certified Copy of Priority Document(s)
11. ☒ Return Receipt Postcard
12. ☐ Other:

**Filing fee calculation:**

- ☐ Before calculating the filing fee, please enter the enclosed Preliminary Amendment.  
☐ Before calculating the filing fee, please cancel claims .

Basic Filing Fee								\$	710
Multiple Dependent Claim Fee (\$ 270)								\$	270
Foreign Language Surcharge (\$ 130)								\$	
	For	Number Filed		Number Extra		Rate			
Extra Claims	Total Claims	30	-20	10	x	\$	18	= \$ 180	
	Independent Claims	5	-3	2	x	\$	80	= \$ 160	
TOTAL FILING FEE								\$	1,320

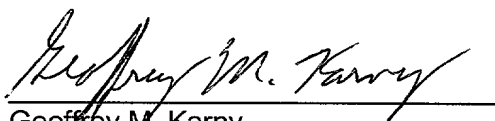
- ☒ Please charge Deposit Account No. 19-0134 in the name of Novartis Corporation in the amount of \$1,320. An additional copy of this paper is enclosed. The Commissioner is hereby authorized to charge any additional fees under 37 CFR §1.16 and §1.17 which may be required in connection with this application, or credit any overpayment, to Deposit Account No. 19-0134 in the name of Novartis Corporation.

Please address all correspondence to the address associated with Customer No. 001095, which is currently:

Thomas Hoxie  
Novartis Corporation  
Patent and Trademark Dept.  
564 Morris Avenue  
Summit, NJ 07901-1027

Please direct all telephone calls to the undersigned at the number given below, and all telefaxes to (301) 258-4620.

Respectfully submitted,



Geoffrey M. Karny  
Attorney for Applicants  
Reg. No. 31,382  
Tel. No. (301) 258-4640

Date: November 17, 2000

## SELECTABLE CELL SURFACE MARKER GENES

This application claims the benefit under 35 USC §119(e) of the following United States provisional patent applications: (1) Provisional Application No. 60/166,594, filed November 19, 1999, for "Selectable Cell Surface Marker Genes;" (2) Provisional Application No. to be assigned, filed November 19, 1999, as Application No. 09/444,038 for "Selective Marker Genes," and subject to a Petition for Conversion to Provisional Application, filed November 16, 2000; and (3) Provisional Application No. to be assigned, filed March 30, 2000 as Application No. 09/539,248 for "Selectable Marker Genes," and subject to a Petition for Conversion to Provisional Application, filed November 16, 2000. The disclosures of these three provisional applications are incorporated herein by reference in their entirety.

This invention relates to a method of identifying genetically modified cells using a mutated protein-tyrosine kinase receptor (PTKR), particularly a mutated epidermal growth factor receptor (EGFR) family member or a mutated muscle specific kinase (MuSK) family member as a selectable cell marker.

### **BACKGROUND OF THE INVENTION**

The use of selectable markers is well known for the identification of prokaryotic and eukaryotic cells, and the use is essential because frequently when a DNA sequence of interest is introduced into a cell it will not necessarily lead to a phenotype that is readily determined. The number of selectable markers used in identifying eukaryotic cells and especially mammalian cells has been limited. In the past, selectable markers that conferred drug resistance were employed (i.e. G-418 and hygromycin). More recently, selectable markers that are combined with fluorescence activated cell sorting (FACS) have been used, for example, green fluorescent protein (GFP). Alternatively, antibodies that recognize a cell surface molecule may be coupled to a fluorophore to help identify the cells of interest.

Several cell surface molecules have been used as a selectable cell marker including murine CD8, CD24, and human Low-Affinity Nerve Growth Factor Receptor (NGFR). Reference is made to the following publications; WO95/06723; WO98/19540; Jolly, et al., Proc. Natl. Acad.

Sci. 80: 477 (1983); and Reddy, et al., Mol. Brain Res. 8:137 (1990). Some of these cell surface molecules have been mutated in their intracellular domain to avoid signaling of the molecule when binding to their ligand. However upon ligand binding some of the intracellularly mutated molecules may homo-, heterodimerize or trimerize. If a newly introduced molecule in the cell should heterodimerize with endogenous receptors a dominant negative effect may result. The present invention provides cell surface markers that do not heterodimerize with endogenous receptors.

### **SUMMARY OF THE INVENTION**

The present invention provides a method of identifying genetically modified mammalian cells comprising introducing a nucleic acid sequence encoding a mutated protein-tyrosine kinase receptor (PTKR) operatively linked to an expression control sequence into a cell to form a genetically modified cell; allowing expression of the mutated PTKR in the genetically modified cell; and identifying said genetically modified cell expressing the mutated PTKR. In a preferred embodiment the mutated PTKR is a mutated epidermal growth factor receptor (EGFR) family member (or MuSK family member), particularly preferred is a mutated EGFR1 and more specifically the sequences given the designation EGFR1-I and EGFR1-II. In a second embodiment the introducing step is accomplished by incorporating the nucleic acid sequence encoding the mutated PTKR, and particularly the mutated EGFR into a vector and introducing the vector into a cell. The preferred vectors are retroviral vectors. In a further embodiment the mammalian cells are human cells, particularly hematopoietic cells, liver cells, endothelial vascular cells, and smooth muscle cells, more particularly hematopoietic cells. In yet another embodiment a heterologous gene is incorporated into the construct or vector comprising the mutated PTKR marker sequence. In yet a further embodiment the identifying step is accomplished by contacting the genetically modified cells with an antibody that recognizes and binds to the mutated PTKR. In yet another embodiment the identifying step separates the genetically modified cells from the non-genetically modified cells.

In a second aspect, the invention provides a method of identifying genetically modified mammalian cells comprising the steps of incorporating into a vector a nucleic acid sequence

encoding a mutated epidermal growth factor receptor (EGFR) family member operatively linked to an expression control sequence; introducing the vector into a mammalian cell to form a genetically modified cell; allowing expression of the mutated EGFR in the genetically modified cell; and identifying said genetically modified cell expressing the mutated EGFR. In a preferred embodiment the mutated EGFR is EGFR1-I or EGFR1-II and the vector is a retroviral vector.

In a third aspect, the invention is directed to a method for the immunoselection of transduced mammalian cells comprising, retrovirally transducing mammalian cells with a nucleic acid sequence encoding a mutated epidermal growth factor receptor (EGFR) family member operatively linked to an expression control sequence; incubating the transduced cells with a marked antibody that recognizes and binds specifically to the mutated EGFR; and identifying the marked transduced cells. In a preferred embodiment the cells are human cells, particularly hematopoietic cells. In another embodiment the cells are transduced by a retroviral vector derived from the group consisting of moloney murine leukemia virus (MoMLV), myeloproliferative sarcoma virus (MPSV), murine embryonic stem cell virus (MESV), murine stem cell virus (MSCV) and spleen focus forming virus (SFFV). In a further embodiment the cells are transduced with a vector derived from a lentivirus. In yet a further embodiment the method includes the step of separating the identified marked transduced cells from non-marked cells. In yet another embodiment the method includes the step of expanding the marked transduced cells.

In a fourth aspect, the invention is directed to a method of identifying mammalian cells expressing a protein of interest, comprising the steps of, introducing into a mammalian cell a nucleic acid sequence encoding a mutated PTKR operatively linked to an expression control sequence and a nucleic acid sequence encoding a protein of interest; culturing the resulting mammalian cells; and identifying cells which express the mutated PTKR thereby obtaining cells which express the protein of interest.

## **BRIEF DESCRIPTION OF THE DRAWINGS**

Figure 1A is a schematic representation of a wild-type (WT) EGFR molecule and two  
 5 mutated EGFRs: EGFR1-I with amino acid residues deleted from the cytoplasmic domain and  
 EGFR1-II with amino acid residues deleted from both the cytoplasmic and extracellular  
 domains. Fig. 1B illustrates the wild-type and the mutated MuSK.

Figure 2 illustrates the amino acid sequence of a wild-type (WT) EGFR1 and  
 corresponds to SEQ ID NO: 2. The signal sequence is represented as amino acid residue 1 to 24.  
 10 The extracellular domain includes amino acid residues 25 through 645, the transmembrane  
 domain includes amino acid residues 646 through 668, and the cytoplasmic domain includes  
 amino acid residues 669 through 1210. The tyrosine kinase domain is located at amino acid  
 residues 718 through 964, and the threonine phosphorylation site is located at amino acid residue  
 678.

Figure 3 illustrates the nucleotide sequence of the wild-type (WT) EGFR1 encoding the  
 15 amino acid sequence of Figure 2 and corresponds to SEQ ID NO: 1.

Figure 4 illustrates the general scheme used to generate a deletion in the extracellular  
 domain and intracellular domain of the EGFR sequence depicted in Figure 2 and 3.

Figure 5 illustrates expression of EGFR1-II on primary T cells and CD34<sup>+</sup> cells after  
 20 transduction with supernatants that were generated from PPA-6 cells. Panel A corresponds to  
 untransduced cells, panel B illustrates expression after transduction and panel C illustrates  
 expression after immuno-magnetic beadselection for mutated EGFR expression on T-cells.  
 Panel D corresponds to untransduced cells and panel E illustrates expression of CD34<sup>+</sup> cells.

Figure 6 illustrates a MuSK-R designated hMuSK-R and corresponds to the nucleic acid  
 25 sequence as set forth in SEQ ID NO:7 and the amino acid sequence as set forth in SEQ ID NO:  
 8.

## **DETAILED DESCRIPTION OF THE INVENTION**

30 The practice of the present invention will employ, unless otherwise indicated  
 conventional techniques of cell biology, molecular biology, cell culture, immunology, virology,

and the like which are in the skill of one in the art. These techniques are fully disclosed in the current literature and reference is made specifically to Sambrook, Fritsch and Maniatis eds., "Molecular Cloning, A Laboratory Manual", 2<sup>nd</sup> Ed., Cold Springs Harbor Laboratory Press, (1989); Celis, J.E. "Cell Biology, A Laboratory Handbook", Academic Press, Inc. (1994);  
 5 Coligan et al., "Current Protocols in Immunology", John Wiley and Sons (1991); and Harlow et al., "Antibodies: A Laboratory Manual" (1988), Biosupplynet Source Book (1999), Cold Springs Harbor Laboratory.

All publications and patent applications cited in this specification are indicative of the level of skill of those skilled in the art to which this invention pertains, and are hereby  
 10 incorporated by reference in their entirety.

As used in this specification and the claims, the singular form "a", "an" and "the" include plural references unless the context clearly dictates otherwise. For example, a stem cell includes a plurality of stem cells.

15 The selectable marker of the instant invention is a mutated protein-tyrosine kinase receptor (PTKR) molecule. PTKR molecules are activated by polypeptide ligands, and are closely related in their catalytic domains. They are Type I transmembrane proteins, with their N-termini outside the cell and single membrane-spanning regions. (See van der Geero et al., Annu. Rev. Cell Biol. 10:251 (1994)). As used herein PTKRs include but are not limited to the  
 20 following subfamilies, MuSK-R which is believed to initiate the formation of neuromuscular junctions in response to agrin (Glass, et al. *Cell* 85:513 (1996), epidermal growth factor receptor (EGFR), platelet-derived growth factor receptor (PDGFR), insulin receptor (INSR), nerve growth factor receptor (NGFR), fibroblast growth factor receptor (FGFR), and EPH. Each subfamily is comprised of various members. The PDGFR family includes PDGF receptor  $\alpha$ ,  
 25 PDGF receptor  $\beta$ , SCF receptor (c-Kit), CSF-1 receptor (c-Fms), Flk2 (Flt3), FLT1 (Quek1), Flk1 (KDR), and FLT4 (Quk2). The INSF subfamily includes insulin receptor, IGF-1 receptor, IRR, ROS and Ltk. The NGFR subfamily includes NGFR receptor (TrkA), TrkB, and TrkC. The EGFR family includes EGFR1, EGFR2, EGFR3, and EGFR4. One skilled in the art is aware of alternative nomenclature for these members e.g. EGFR1 has been referred to as EGFR,  
 30 HER, c-ErbB, and ErbB-1; EGFR2 has been referred to as HER2, Neu and ErbB-2; EGFR3 has

been referred to as HER3 and ErbB-3; and EGFR4 has been referred to as HER4 and ErbB-4. (Ullrich and Schlessinger, Cell, 61: 203-212 (1990)).

The mutated PTKRs of the present invention include a modification in a PTKR molecule so that the receptor no longer possesses the signaling activity of the corresponding unmodified (wild type) receptor. For the purpose of the present disclosure, "Wild-Type PTKRs" include not only naturally occurring PTKRs but also may include genetically engineered PTKRs molecules wherein the PTKR molecule has undergone changes in the DNA sequence that do not significantly effect the properties of the protein tyrosine kinase receptor molecule. These changes include ones that do not change the encoded amino acid sequence, ones that result in conservative substitutions of amino acid sequences, or ones that result in one or a few amino acid deletions or additions. Suitable substitutions are known by those skilled in the art. Amino acid residues which can be conservatively substituted for one another include, but are not limited to, glycine/alanine; valine/isoleucine/leucine, asparagine/glutamine; aspartic acid/glutamic acid; serine/threonine; lysine/arginine; and phenylalanine/tyrosine.

A modification of the wild-type PTKR molecule may include deletions, truncations and the like. More specifically, mutated PTKRs of the invention include a modification, however, obtained, in the cytoplasmic domain causing the molecule to be devoid of signaling activity. Signaling activity may be generally defined as activity that triggers a response pathway in the cytosol to the nucleus that ultimately leads to activation of transcription. The mutated PTKRs according to the invention may include a modification in the extracellular domain. However, preferably the extracellular domain should still be capable of binding an antibody. In general, the smallest peptide fragment of the extracellular domain capable of binding an antibody would be approximately 15 amino acid residues, more preferably at least 50 amino acid residues.

In a preferred embodiment, the selective marker PTKR is a mutated member of the EGFR family. The members of the EGFR family have similar overall topology and significant relatedness in amino acid sequence. Members of this family are not limited to the above-enumerated members, but include new members sharing common features, such as a single transmembrane domain; two cysteine rich regions of approximately 100 to 200 amino acid residues, and a single tyrosine kinase domain. Further, members of the EGFR family can heterodimerize with other members of the EGFR family.



A preferred mutated EGFR family member is EGFR1. EGFR1 is a high affinity receptor for a number of ligands including but not limited to epidermal growth factor (EGF), transforming growth factor  $\alpha$  (TGF $\alpha$ ) and amphiregulin. Various EGFR sequences have been identified and reference is made to EMBL/GenBank Accession Numbers: X00588/X06370 (Ullrich, et al., Nature 309:418 (1984)); K01885/K02047, (Lin et al., Science 224:843 (1984)); K03193 (Merlino, et al., Mol. Cell. Biol 5:1722 (1985)); X00663 (Xu, et al., Nature 309:806 (1984); AC006977; M38425 (Simmen, et al., Biochem. Biophys. Res. Commun. 124:125 (1984)); and M20386 (Lax, et al., Mol. Cell. Biol. 8:1970 (1988)). Several EGFRs have been identified in invertebrate species, for example *Drosophila* having EMBL/GenBank Accession Number K03054 (Livneh, et al., Cell 40:599 (1985)).

The domain structure of wild-type EGFR is schematically illustrated in Figure 1. EGFR is comprised of a signal peptide that targets the protein to the secretory pathway. The extracellular domain follows the signal sequence. This domain is made up of several hundred amino acids. The extracellular domain means the part of the receptor that normally projects from the cell into the extracellular environment and in the EGFR it contains distinctive Cys residues. The transmembrane domain is generally localized in the cell membrane and consists of a stretch of hydrophobic residues followed by several basic residues. The cytoplasmic domain (also referred to as the intracellular domain) includes the catalytic part of the molecule and is positioned within the cell. This domain contains the substrate binding site and regulatory tyrosine and threonine phosphorylation sites.

Modifications of EGFR family members are known, and particularly mutations of EGFR1 are known. WO 93/05148 discloses three mutations of EGFR1. One mutation (referred to as HER721A) is a point mutation in the cytoplasmic domain of EGFR1 at amino acid position 721. The lysine is changed to alanine. A second mutation (referred to as HERCD-533) involves the C-terminal end of the molecule wherein the intracellular domain was deleted but the transmembrane domain was not. A third mutation (referred to as HERCD-566) includes the C-terminal deletion of 566 amino acids, both the intracellular and transmembrane domains were deleted. Three truncated forms of EGFR have been detected in some malignancies: Type I (EGFRvIII) has a large deletion of the extracellular domain (amino acids 6 – 273) and does not bind EGF (Moscattello, D.K. et al., Oncogene 13:85 (1996)). Type II contains an in-frame deletion of 83 amino acids (520–603) in extracellular domain IV that does not prevent EGF and

TGF $\alpha$  binding (Humphrey, P.A. et al., Biophys. Res. Commun. 178:1413 (1991)). Type III has an in-frame deletion of 267 amino acids (29-296) in extracellular domains II and III. These two mutations prevent ligand binding (Humphrey, P.A. et al., Proc. Natl. Acad. Sci. USA 87:4207 (1990)).

While specific modifications such as, but not limited to those enumerated above may be known for the EGFR members, the use of a mutated EGFR family member as a selectable marker in mammalian cells is not known. According to the invention preferred modifications to the PTKR molecules and particularly to EGFR molecules include modifications to the cytoplasmic domain, such as deletions of at least 150, preferably at least 250, more preferably at least 400, and still more preferably at least 500 amino acids of the cytoplasmic domain. In a preferred embodiment, the deletions will consist of a truncation of the molecule. Truncations may include deletion of tyrosine phosphorylation sites in the range of 1 to 15 sites and deletion of the kinase catalytic site. However, a mutated EGFR useful as selectable marker in the invention may include more than 1 and fewer than 15 deletions of tyrosine phosphorylation sites. Additionally, the above modifications may include in-frame deletions, such as deletions of at least 50 amino acids of the cytoplasmic domain. These deletions may include deletion of the protein tyrosine-kinase activity. Particularly preferred selectable marker sequences are mutated EGFR1, EGFR2 and EGFR3 molecules wherein at least 400 and preferably 500 amino acids of the cytoplasmic domain are deleted from the corresponding wild-type molecule.

In the present invention, a preferred mutated EGFR molecule is a mutated EGFR1. Particularly preferred is a modified sequence derived from EGFR1 as illustrated in Figures 2 and 3 (SEQ ID Nos. 1 and 2). In this EGFR1 molecule, the extracellular domain is encoded by nucleotides 1 through 1935, the transmembrane domain is encoded by nucleotides 1938 through 2004, and the intracellular domain is coded by nucleotides 2007 to 3630. One non-limiting example of a preferred mutated EGFR1 marker is designated EGFR1-I wherein amino acid sequence 679 to 1210 of the cytoplasmic domain as illustrated in Figure 2 is deleted.

In a second preferred embodiment, the selective marker PTKR is a mutated member of the MuSK-R family. MuSK-R is comprised of a signal sequence or leader sequence that targets the protein to the secretory pathway. The extracellular domain follows the signal sequence. This domain is made up of several hundred amino acids, and while the exact number of amino acid residues varies, typically the extracellular domain includes around 500 amino acids. The

extracellular domain is the part of the receptor that normally projects from the cell into the extracellular environment and includes a ligand binding region. The extracellular domain is one of the most distinctive features of the kinase receptors. In MuSK-R, the extracellular domain contains immunoglobulin-like (Ig-like) regions. Typically four Ig-like regions are found. However there are reports of MuSK-Rs with three Ig-like regions. The extracellular domain may include 6 contiguous cysteine residues known as a C6-box. While the location of the C6-box may vary depending on the particular MuSK-R, in certain MuSK-Rs it is found approximately at amino acid residues 373 - 382. The transmembrane domain is generally localized in the cell membrane and consists of a stretch of hydrophobic residues followed by several basic residues. The intracellular domain (used interchangeably with the cytoplasmic domain) includes the catalytic part of the molecule and is positioned within the cell.

MuSK-Rs are also known in the art as denervated muscle kinase receptors and have been referred to as DmKs (see U.S. Pat. No. 5,656,473 and particularly SEQ ID NOS: 16 and 17 therein). MuSK-R sequences have been isolated and identified from humans, rats, mice, and xenopus. Closely related to human MuSK-R is a receptor isolated from the electric ray *Torpedo californica* designated Torpedo tyrosine kinase receptor, and ROR tyrosine kinase receptors (Jennings, et al. *Proc. Natl. Acad. Sci. USA* 90:2895 (1993); Masiakowski et al., *J. Biol. Chem.* 267: 26181 – 26190 (1992); Valenzuela et al., *Neuron*, 15:573-584 (1995); and Hesser et al., *FEBS Letters*, 442:133- 137 (1999)).

Other non-limiting examples of MuSK-Rs available from public depositories such as GeneBank and ATCC include accession numbers: NM005592; AF006464; A448972; AI800924; AI700028; AI341265; AI341122; AI302067; U34985; AA448972; and ATCC 75498. As mentioned above MuSK-R is specific to the skeletal muscle lineage.

The term MuSK-R as used in the present specification and claims is broadly defined to include the known MuSK-Rs (including DmK receptors), isoforms or variants of known MuSK-Rs having similar structure, tyrosine kinase receptors that are functionally similar to known MuSK-Rs and novel MuSK-Rs not previously described that are identified using screening techniques well known to those in the art. Such techniques may include the use of degenerate oligodeoxyribonucleotide primers.

Accordingly, the term MuSK-R when referring to a nucleic acid molecule includes (a) nucleic acid sequences comprising a coding region of a known mammalian MuSK-R; (b) a

nucleic acid sequence which hybridizes under stringent conditions to the nucleic acid of (a) and which encodes a mammalian MuSK-R; and (c) a degenerate MuSK-R wherein the MuSK-R has undergone changes in its nucleic acid sequence that does not significantly effect the properties of the MuSK-R protein encoded by the polynucleotide. These changes include ones that do not  
 5 change the encoded amino acid sequence, ones that result in conservative substitutions of amino acid sequences, or ones that result in one or a few amino acid deletions or additions. Suitable substitutions are known by those skilled in the art. Amino acid residues that can be conservatively substituted for one another include but are not limited to, glycine/alanine; valine/isoleucine/leucine, asparagine/glutamine; aspartic acid/glutamic acid; serine/threonine;  
 10 lysine/arginine; and phenylalanine/tyrosine. Any conservative amino acid substitution not significantly affecting the properties of a MuSK-R is encompassed by the term. MuSK-R includes not only naturally occurring MuSK-Rs but also may include genetically engineered MuSK-Rs.

The term MuSK-R when referring to a polypeptide encompasses known MuSK  
 15 receptors, isoforms or variants of MuSK-Rs, and functionally equivalent receptors. A functionally equivalent receptor is a MuSK-R that can compete with a known MuSK-R for binding. More specifically, a functionally equivalent MuSK-R has at least 40%, preferably at least 60%, and more preferably at least 80% identical amino acids to the sequence set forth in SEQ ID NO: 8 and can compete with the MuSK-R illustrated in figure 6 for ligand or substrate  
 20 binding.

According to the invention mutated MuSK-R are used as selective markers to identify genetically modified cells. The marker is introduced on a nucleic acid construct into a target cell that normally does not express a MuSK-R. The term "introduced" is broadly used herein to include inserted, incorporated and the like. When the mutated MuSK-Rs are used as selectable  
 25 markers the molecule no longer possesses signaling activity. Signaling activity has been generally defined as triggering a response pathway in the cytosol to the nucleus which ultimately leads to activation of transcription. The lack of signaling activity may be due to a) use of a MuSK-R in tissue or cells other than muscle (Glass et al., *Cell* 85:513-523 (1996)) or b) use of a mutated MuSK-R.

30 While modifications of MuSK-R may be known, the method of identifying genetically modified cells comprising using a mutated MuSK-R as a selectable marker is not known.

As stated above, the localization of MuSK-R is in muscle tissue and MuSK-R serves as the functional agrin receptor. Agrin is a factor that can induce molecular reorganizations at the motor endplate. Therefore, MuSK-R may be used as a selective marker in tissue other than muscle.

5 In a preferred embodiment, the modifications to a MuSK-R include truncations and/or deletions of MuSK-Rs. The mutation may occur in the extracellular domain and/or the intracellular domain by means well known in the art. The mutation causes the molecule to be devoid of signaling activity. Preferably the extracellular domain should still be capable of binding an antibody. In general the smallest peptide fragment of the extracellular domain  
10 capable of binding an antibody would be approximately 15 amino acid residues, more preferably at least 50 amino acid residues.

A preferred MuSK-R according to the invention is the sequence set forth in SEQ ID NOs: 7 and 8, designated herein as hMuSK-R. The extracellular domain is encoded by nucleotides 1 through 1479, the transmembrane domain is encoded by nucleotides 1480 through  
15 1545, and the intracellular domain is encoded by nucleotides 1546 through 2607. Other preferred MuSK-Rs are molecules closely related to the sequences set forth in SEQ ID NO: 8 and 7. Examples of closely related sequences are the sequences set forth in U.S. Pat No. 5,656,473 particularly NO: 16 and 17 disclosed therein.

Mutants of MuSK-R are known and reference is made to Apel et al., *Neuron* 18:623 –  
20 635 (1997). In the present invention, preferred modifications to a MuSK-R include modifications to the cytoplasmic domain such as deletions of at least 150, preferably at least 200, more preferably at least 250, more preferably 300, and most preferably at least 350 amino acids of the cytoplasmic domain. The deletions are preferably truncations. Deletions or truncations may include deletion of tyrosine phosphorylation sites in the range of 1 to 19,  
25 preferably 2 – 15, more preferably 2 – 10 sites. Additionally the kinase catalytic site may be deleted from a MuSK-R. In one aspect, the kinase catalytic site is found at approximately amino acid residues 672 to 691 of SEQ ID NO: 8. As long as the protein is stably expressed, there is no limitation to the number of sequences deleted or truncated in the cytoplasmic domain.

Particularly preferred mutated MuSK-Rs useful as selectable markers according to the  
30 invention include modifications to the MuSK-R sequence set forth in Figure 6 (SEQ ID NO: 8). In one embodiment the MuSK-R is truncated by least 300 amino acid residues in the

cytoplasmic domain. One preferred embodiment includes the deletion of amino acid sequence 538 – 869 and is designated mMuSK-RI. Another preferred embodiment includes the deletion of amino acid sequence 577 – 869 and is designated mMuSK-RII.

In addition to modification of the cytoplasmic domain mutations may be made in the extracellular domain. The extracellular domain modification may include deletion of at least about 100 amino acids, preferably at least about 150 amino acids, more preferably at least about 200 amino acids, and still more preferably at least about 250 amino acids. A mutated MuSK-R used as a selectable marker according to the invention preferably should contain an antibody-binding site in the extracellular domain.

In another preferred embodiment, a mutated PTKR, particularly a mutated EGFR or mutated MusSK-R useful as a selectable marker includes both modification of the cytoplasmic domain such as described above, and modification to the extracellular domain. This is expected to avoid heterodimerization of the mutated PTKR with endogenous receptors.

The extracellular domain modifications may include deletion of at least about 100 amino acids, preferably at least about 150 amino acids, more preferably at least about 200 amino acids, and still more preferably at least about 250 amino acids. In one embodiment, the modification will include removal of one cysteine-rich area generally characterized by an area of about 100 amino acids. Particularly preferred marker sequences are mutated EGFR1, EGFR2 and EGFR3 molecules having not only a truncation in the cytoplasmic domain but also a deletion of at least 200 and preferably at least 250 amino acids of the extracellular domain. As mentioned above, a mutated PTKR as a selectable marker useful in the methods of the invention should preferably include an antibody-binding site in the extracellular domain.

More particularly, a preferred mutated EGFR molecule wherein both the extracellular domain and the cytoplasmic domain are modified is a sequence derived from WT EGFR1 illustrated in Figure 2 (SEQ ID Nos. 1 and 2). Particularly preferred is a mutated EGFR1 sequence designated EGFR1-II wherein amino acid sequence 25 through 313 of the extracellular domain and amino acid sequence 679 through 1210 of the cytoplasmic domain of Figure 2 (SEQ ID NO. 2) are deleted.

As used herein a “mutated PTKR” or a “mutated EGFR” or a “mutated MuSK-R” selectable marker refers to nucleotides or protein as appropriate from context. Polynucleotides

of the invention may be in the form of RNA or in the form of DNA, which DNA includes cDNA, genomic DNA or synthetic DNA.

General strategies for creating mutations in nucleic acids and proteins are well known. These methods may be used to create mutant PTKRs, particularly mutant EGFR family members useful as selectable markers in the present invention. One skilled in the art is aware that PTKR molecules including members of the EGFR family may be obtained from sources such as various sequence databases including GenBank. Both random and site-directed mutagenesis methods may be effective to create mutations in wild type PTKRs. Random methods encompass altering the sequences within restriction endonuclease sites, inserting an oligonucleotide linker into a plasmid, using chemicals to damage plasmid DNA, and incorporating incorrect nucleotides during in vitro DNA synthesis. However, site-directed mutagenesis may be a more beneficial tool. Particularly preferred site-directed methods include oligonucleotide-directed mutagenesis and polymerase chain reaction (PCR) –amplified oligonucleotide-directed mutagenesis. These methods are also well known in the art and reference is made to Wu, et al., eds. *Methods in Enzymology*, Vol. 154: Recombinant DNA, Part E, Academic, NY (1987); Landt, et al., *Gene*, 96:125 – 128 (1990); Kirchhoff et al., *Methods Mol. Biol.* 57:323–333 (1995); Herlitz, et al., *Gene* 91:143–147 (1990); Sambrook, et al. *Molecular Cloning, A Laboratory Manual*, *supra*, and *Current Protocols in Molecular Biology*, Greene Publishing Associates, John Wiley & Sons, Inc Canada (1988).

The usefulness of a mutated PTKR as a selectable marker concerns the ability to identify genetically modified mammalian cells *in vitro*, *ex vivo*, and *in vivo*. While the mutated PTKR sequence may be introduced into a target cell as part of a nucleic acid construct operatively linked to an expression control sequence, in a preferred embodiment the construct including the mutated PTKR sequence is placed in a vector and then introduced into a target cell. As used herein “operatively linked” refers to an arrangement of elements wherein the components are configured so as to perform their usual function. The control elements need not be contiguous with the coding sequence.

Vectors containing both a promoter and a cloning site into which a polynucleotide can be operatively linked are well known in the art. Such vectors are capable of transcribing RNA *in vitro* or *in vivo*, and are commercially available from sources such as Stratagene (La Jolla, CA) and Promega Biotech (Madison, WI). Examples of vectors include vectors derived from viruses,

such as baculovirus, retroviruses, adenoviruses, adeno-associated viruses, and herpes simplex viruses; bacteriophages; cosmids; plasmid vectors; fungal vectors; synthetic vectors; and other recombination vehicles typically used in the art. These vectors have been described for expression in a variety of eukaryotic and prokaryotic hosts and may be used for protein expression.

In a preferred embodiment, the vector comprises a nucleic acid sequence coding for a selective marker according to the invention, operatively linked to an expression control sequence. Selection of appropriate control sequences is dependent on the target cell used and the choice is within the skill of one in the art. Examples of expression control sequences, also referred to as regulatory sequences, include promoters, enhancers, polyadenylation signals, RNA polymerase binding sequences, sequences conferring inducibility of transcription and other expression control elements, such as scaffold attachment regions (SARs).

The promoter may be either a prokaryotic or eukaryotic promoter. Additionally the promoter may be a tissue specific promoter, inducible promoter, synthetic promoter, or hybrid promoter. More than one promoter may be placed in the construct. Examples of promoters include but are not limited to the phage lamda (PL) promoter; SV40 early promoter; adenovirus promoters, such as adenovirus major late promoter (Ad MLP); herpes simplex virus (HSV) promoter; a cytomegalovirus (CMV) promoter; such as the human CMV immediate early promoter; a long terminal repeat (LTR) promoter, such as a MoMLV LTR; the U3 region promoter of the Moloney murine sarcoma virus; Granzyme A promoter; regulatory sequences of the metallothionein gene; CD34 promoter; CD8 promoter; thymidine kinase (TK) promoters; B19 parvovirus promoters; and rous sarcoma virus (RSV) promoter. Additionally promoter elements from yeast and other fungi may be used, such as Gal 4 promoter and the alcohol dehydrogenase (ADH) promoter. These promoters are available commercially from various sources such as Stratagene (La Jolla, CA). It is to be understood that the scope of the present invention is not to be limited to a specific promoter.

The vector may further comprise a polyadenylation signal that is positioned 3' of the carboxy-terminal amino acid. Vectors containing both a promoter and a cloning site into which a polynucleotide can be operably linked are well known in the art. Such vectors are capable of transcribing RNA *in vitro* or *in vivo*, and are commercially available. Specific non-limiting examples include pSG, pSV2CAT, and pXt1 from Stratagene (La Jolla, CA) and pMSG, pSVL,



pBPV and pSVK3 from Pharmacia. Other exemplary vectors include the pCMV mammalian expression vectors, such as pCMV6b and pCMV6c (Chiron Corporation, CA), pSFFV-Neo, and pBluescript-SK+. In order to optimize expression and/or in vitro transcription, it may be necessary to remove, add or alter 5' and/or 3' untranslated portions of polynucleotides to eliminate potentially extra inappropriate alternative translation initiation codons or other sequences that may interfere with or reduce expression, either at the level of transcription or translation. Alternatively consensus ribosome binding sites can be inserted immediately '5' of the start codon to enhance expression.

Particularly preferred vectors are retroviral vectors and reference is made to Coffin et al., "Retroviruses", (1997) Chapter 9 pp; 437-473 Cold Springs Harbor Laboratory Press. Retroviral vectors useful in the invention are produced recombinantly by procedures already taught in the art. WO94/29438, WO97/21824 and WO97/21825 describe the construction of retroviral packaging plasmids and packing cell lines. Common retroviral vectors are those derived from murine, avian or primate retroviruses. The most common retroviral vectors are those based on the Moloney murine leukemia virus (MoMLV) and mouse stem cell virus (MSCV). Vectors derived from MoMLV include, LMily, LINGFER, MINGFR, MND and MINT (Bender, et al., J. Virol. 61:1639-1649 (1987); Miller, et al., Biotechniques 7: 98-990 (1989); Robbins, et al., J. Virol. 71:9466-9474 (1997) and U. S. Pat. No. 5,707,865). Vectors derived from MSCV include MSCV-MiLy (Agarwal, et al., J. of Virology 72:3720). Further non-limiting examples of vectors include those based on Gibbon ape leukemia virus (GALV), Moloney murine sarcoma virus (MoMSV), myeloproliferative sarcoma virus (MPSV), murine embryonic stem cell virus (MESV), spleen focus forming virus (SFFV) and the lentiviruses, such as Human immunodeficiency virus (HIV-1 and HIV-2). New vector systems are continually being developed to take advantage of particular properties of parent retroviruses such as host range, usage of alternative cell surface receptors and the like (See C. Baum et al., Chapter 4 in Gene Therapy of Cancer Cells eds., Lattime and Gerson (1998)). The present invention is not limited to particular retroviral vectors, but may include any retroviral vector. Particularly preferred vectors include DNA from a murine virus corresponding to two long terminal repeats, and a packaging signal. In one embodiment the vector is a MoMLV or MSCV derived vector. In another preferred embodiment the vector is MND.

In producing retroviral vector constructs, the viral gag, pol and env sequence will generally be removed from the virus, creating room for insertion of foreign DNA sequences. Genes encoded by foreign DNA are usually expressed under the control a strong viral promoter in the long terminal repeat (LTR). While a LTR promoter is preferred, as mentioned above, numerous promoters are known.

Such a construct can be packaged into viral particles efficiently if the gag, pol and env functions are provided in trans by a packaging cell line. Therefore when the vector construct is introduced into the packaging cell, the gag-pol and env proteins produced by the cell, assemble with the vector RNA to produce infectious virions that are secreted into the culture medium. The virus thus produced can infect and integrate into the DNA of the target cell, but does not produce infectious viral particles since it is lacking essential packaging sequences. Most of the packaging cell lines currently in use have been transfected with separate plasmids, each containing one of the necessary coding sequences, so that multiple recombination events are necessary before a replication competent virus can be produced. Alternatively the packaging cell line harbors a provirus. (The DNA form of the reverse-transcribed RNA once its integrates into the genomic DNA of the infected cell). The provirus has been crippled so that although it may produce all the proteins required to assemble infectious viruses, its own RNA can not be packaged into virus. RNA produced from the recombinant virus is packaged instead. Therefore, the virus stock released from the packaging cells contains only recombinant virus. Non-limiting examples of retroviral packaging lines include PA12, PA317, PE501, PG13, ΨCRIP, RD114, GP7C-tTA-G10, ProPak-A (PPA-6), and PT67. Reference is made to Miller et al., *Mol. Cell Biol.* 6:2895 (1986); Miller, et al., *Biotechniques* 7:980 (1989); Danos, et al., *Proc. Natl. Acad. Sci. USA* 85:6460 (1988); Pear, et al., *Proc. Natl. Acad. Sci. USA* 90:8392 (1993); Rigg, et al., *Virology* 218: (1996); and Finer et al., *Blood* 83:43 (1994). Retroviral vector DNA can be introduced into packaging cells either by stable or transient transfection to produce vector particles.

Additionally preferred vectors include adenoviral vectors (See Frey et al., *Blood* 91:2781 (1998) and WO95/27071) and adeno-associated viral vectors (AAV) (See Chatterjee et al., *Current Topics in Microbiol. and Immunol.* 218:61 (1996). Reference is also made to Shenk, Chapter 6, 161– 78, Breakefield et al., Chapter 8 201-235; Kroner-Lux et al., Chapter 9, 235–256 in *Stem Cell Biology and Gene Therapy*, eds. Quesenberry et al., John Wiley & Sons,

1998 and U.S. Pat. Nos. 5,693,531 and 5,691,176. The use of adenovirus derived vectors may be advantageous under certain situations because they are capable of infecting non-dividing cells, and unlike retroviral DNA, the adenoviral DNA is not integrated into the genome of the target cell. Further the capacity to carry foreign DNA is much larger in adenoviral vectors than retroviral vectors. The adeno-associated viral vectors are another useful delivery system. The DNA of these viruses may be integrated into non-dividing cells, and a number of polynucleotides have been successfully introduced into different cell types using adeno-associated viral vectors. The vectors are capable of transducing several cell types including hematopoietic cells and epithelial cells.

Vectors may also include hybrid vectors of AAV and adenoviruses as described in WO96/13598 and WO99/47691 (The Trustees of the University of Pennsylvania), WO98/21345 (General Hospital), US5965441 (General Hospital), or WO99/58700 (Ariad Gne Therap.), the teaching of which being incorporated into the present invention in their entirety.

In one embodiment, the construct or vector will include not only a nucleic acid sequence encoding a mutated PTKR or a mutated EGFR as a selectable marker but also a second nucleic acid sequence encoding a heterologous gene to be transferred into a target cell. In a preferred embodiment the nucleic acid molecules are DNA. The term "heterologous" as it relates to nucleic acid sequences such as gene sequences denotes sequences that are not normally associated with a particular cell or vector and which are suitably inserted into a construct or vector under control of a promoter to permit expression in the target cell to be genetically modified. The heterologous gene may be located 5' or 3' to the selectable marker of the invention.

Non-limiting preferred vector constructs comprise the general structure as outlined below 5' to 3':

- (a) LTR-X-I-M-LTR;
- (b) LTR-M-LTR;
- (c) LTR-M-(I)-X-LTR;
- (d) LTR-X-pM-LTR; and
- (e) LTR-X-I-M-SAR-LTR

wherein LTR is a long terminal repeat, X is a heterologous gene for a desired protein, M is a selectable marker of the invention, I is an internal ribosomal entry site, SAR is a scaffold attachment region and p is a second promoter, preferably a CMV or PGK promoter.

A gene or coding sequence or a sequence that encodes a particular protein is a nucleic acid molecule that is transcribed and translated into a polypeptide *in vitro* or *in vivo* when placed under the control of appropriate regulatory sequences. The heterologous gene may be any gene for which expression is desired. Examples of heterologous genes include but are not limited to a therapeutic protein, a structural gene, a ribozyme, or an antisense sequence. The heterologous gene may be the entire protein or only the functionally active fragment thereof. The protein may include for example one that regulates cell differentiation or a therapeutic gene capable of compensating for a deficiency in a patient that arises from a defective endogenous gene. Additionally a therapeutic protein or gene may be one that antagonizes production or function of an infectious agent, antagonizes pathological processes, improves a host's genetic makeup, or facilitates engraftment.

Specific examples of a therapeutic gene or gene sequences are ones effective in the treatment of adenosine deaminase deficiency (ADA); sickle cell anemia; recombinase deficiency; recombinase regulatory gene deficiency; HIV, such as an antisense or trans-dominant REV gene; or a gene carrying a herpes simplex virus thymidine kinase (HSV-tk)). The heterologous gene may encode new antigens or drug resistant genes or may encode a toxin or an apoptosis inducer effective to specifically kill cancerous cells, or a specific suicide gene to hematopoietic cells may be included. Additionally the heterologous gene may be a therapeutic gene that is non-human such as a yeast gene (Seo et al., Proc. Natl. Acad. Sci. 95:9167 (1998)).

The vector or construct may also comprise a second heterologous gene in addition to the first heterologous gene encoding a protein of interest. More than one gene may be necessary for the treatment of a particular disease. Alternatively more than one gene can be delivered using several compatible vectors. Depending on the genetic defect, the therapeutic gene can include regulatory and untranslated sequences. For human patients the therapeutic gene will generally be of human origin although genes of closely related species that exhibit high homology and biologically identical or equivalent function in humans may be used if the gene does not produce an adverse immune reaction in the recipient.

Nucleotide sequences of the heterologous gene encoding a protein of interest will generally be known in the art or can be obtained from various sequence databases such as GenBank. One skilled in the art will readily recognize that any structural gene can be excised as a compatible restriction fragment and placed in a vector in such a manner as to allow proper expression of the structural gene in target cells.

The target cells of the invention are mammalian cells and these include but are not limited to humans, mice, monkeys, chimpanzees, farm animals; such as cattle, sheep, pigs, goats, and horses, sport animals, pets; such as dogs and cats, and other laboratory rodents and animals; such as mice, rats, guinea pigs and the like. Preferably the target cells are human cells. Preferred human cells include liver, hematopoietic, smooth muscle, neural, endothelial vascular cells, tumor cells and epithelial cells. Hematopoietic cells are particularly preferred, and these cells encompass hematopoietic stem cells, erythrocytes, neutrophils, monocytes, platelets, mast cells, eosinophils and basophils, B and T lymphocytes, and NK cells as well as the respective lineage progenitor cells. Hematopoietic stem cells and T-cells are especially preferred. Hematopoietic stem cells (HSC) are defined as a population of hematopoietic cells containing long term multilineage repopulating potential. T-cells are defined as a type of lymphocyte and are thought to develop from hematopoietic stem cells. There are many types of T-cells including cytotoxic T-cells, helper T-cells, inducer T-cells and suppressor T cells.

Methods of obtaining target cells, particularly hematopoietic cells are well known in the art and not repeated herein. Non-limiting sources of hematopoietic cells, including hematopoietic stem cells, are bone marrow, embryonic yolk sac, fetal liver tissue, adult spleen, and blood such as adult peripheral blood and umbilical cord blood. (To et al., Blood 89:2233 (1997)). Bone marrow cells may be obtained from ilium, sternum, tibiae, femora, spine and other bone cavities.

The manner in which target cells may be separated from other cells is not critical to this invention. Various procedures may be employed and include physical separation, magnetic separation using antibody-coated magnetic beads, affinity chromatography, and cytotoxic agents joined to a monoclonal antibody or used in conjunction with a monoclonal antibody. Also included is the use of fluorescence activated cell sorters (FACS) wherein the cells can be separated on the basis of the level of staining of the particular antigens. These techniques are well known to those skilled in the art and are described in various references including U.S.

Patent Nos. 5,061,620; 5,409,821; 5,677,136; and 5,750,397; and Yau et al., *Exp. Hematol.* 18:219-222 (1990).

The order of cell separation is not critical to the invention, and specific cell types may be separated either prior to genetic modification with the mutated PTKR or after genetic modification. Preferably cells are initially separated by a coarse separation followed by using positive and/or negative selection. In humans the surface antigen expression profile of an enriched hematopoietic stem cell population may be identified by  $CD34^{+}Thy-1^{+}Lin^{-}$ . Other nonlimiting enriched phenotypes may include:  $CD2^{-}$ ,  $CD3^{-}$ ,  $CD4^{-}$ ,  $CD8^{-}$ ,  $CD10^{-}$ ,  $CD14^{-}$ ,  $CD15^{-}$ ,  $CD19^{-}$ ,  $CD20^{-}$ ,  $CD33^{-}$ ,  $CD34^{-}$ ,  $CD38^{lo/-}$ ,  $CD45$ ,  $CD59^{+/-}$ ,  $CD71^{-}$ ,  $CDW109^{+}$ , glycophorin $^{-}$ , AC133 $^{+}$ , HLA-DR $^{+/-}$ , and EM $^{+}$ . Lin $^{-}$  refers to a cell population selected on the basis of lack of expression of at least one lineage specific marker, such as, CD2, CD3, CD14, CD15 and CD56. The combination of expression markers used to isolate and define an enriched HSC population may vary depending on various factors and may vary as other express markers become available.

Murine HSCs with similar properties to the human  $CD34^{+}Thy-1^{+}Lin^{-}$  may be identified by  $kit^{+}Thy-1.1^{lo}Lin^{-/lo}Sca-1^{+}$  (KTLS). Other phenotypes are well known. When CD34 expression is combined with selection for Thy-1, a composition comprising approximately fewer than 5% lineage committed cells can be isolated (U.S. Patent No. 5,061,620).

It has been shown that CD3 is expressed on most T cells, and that these cells can express the cell surface antigens CD2, CD4, and CD8 antigens. Also CD45 is a useful T cell marker. The most well known T cell marker is the T cell antigen receptor (TCR). There are presently two defined types of TCRs;  $\alpha$ ,  $\beta$ - TCR and  $\gamma$ ,  $\delta$ - TCR. B cells may be selected, for example, by expression of CD19 and CD20. Myeloid cells may be selected for example, by expression of CD14, CD15 and CD16. NK cells may be selected based on expression of CD56 and CD16. Erythrocytes may be identified by expression of glycophorin A. Neuronal cells may be identified by NCAM and LNGFR (Baldwin et al., *J. Cell Biochem.* 15:502 (1996)). Vascular endothelial cells may be identified by VEGFR2, CD34, P-Selectin, VCAM-1, ELAM-1, and ICAM-1.(Horvathova et al., *Biol. Trace Elem. Res.*, 69: 15-26 (1999)). One skilled in the art is aware of other useful markers for the identification of other target cells.

Once a population containing the target cells are harvested and target cells, particularly hematopoietic cells, are separated, the cells are cultured in a suitable medium comprising a

combination of growth factors that are sufficient to maintain growth. Methods for culturing target cells are well known to those skilled in the art, and these methods are only briefly mentioned herein. Any suitable culture container may be used, and these are readily available from commercial vendors. The seeding level is not critical and will depend on the type of cells used, but in general the seeding level for hematopoietic cells will be at least 10 cells per ml, more usually at least about 100 cells per ml and generally not more than  $10^6$  cells per ml when the cells express CD34.

Various culture media, solid or liquid, can be used and non-limiting examples include DMEM, IMDM, X-vivo 15 and RPMI-1640. These are commercially available from various vendors. The formulations may be supplemented with a variety of different nutrients, growth factors, such as cytokines and the like. The medium can be serum free or supplemented with suitable amounts of serum such as fetal calf serum, autologous serum or plasma. If cells or cellular products are to be used in humans, the medium will preferably be serum free or supplemented with autologous serum or plasma. (Lansdorp et al., J. Exp. Med. 175:1501 (1992) and Petzer, et al. PNAS 93:1470 (1996). Also reference is made to Freshney, R. I., "Culture of Animal Cells, A Manual of Basic Techniques", Wiley-Liss, Inc. (1994)).

Non-limiting examples of compounds which may be used to supplement the culture medium are TPO, FL, KL, IL-1, IL-2, IL-3, IL-6, IL-12, IL-11, stem cell factor, G-CSF, GM-CSF, Stl factor, MCGF, LIF MIP-1 $\alpha$  and EPO. These compounds may be used alone or in any combination, and preferred concentration ranges may be readily determined from the published art. When murine stem cells are cultured, a preferred non-limiting medium includes mIL-3, mIL-6 and mSCF. Other molecules can be added to the culture media, for instance, adhesion molecules, such as fibronectin or RetroNectin™ (commercially produced by Takara Shuzo Co., Otsu Shigi, Japan).

In vitro systems for measurement of mammalian stem cell activity include the long-term culture initiating cell assay (LTCIC) and the cobblestone-area-forming cell (CAFC) assay. (Pettengell, et al., Blood 84:3653 (1994); Breems et al., Leukemia 8:1095 (1994); Reading, et al., Exp. Hem. 22:786 (Abst # 406) (1994); and Ploemacher, et al., Blood 74:2755 (1989)). In the CAFC assay a sparsely plated cell population is simply tested for its ability to form distinct clonal outgrowths (or cobblestone areas) on a stromal cell monolayer over a period of time. This assay gives frequency readouts that correlate with LTCIC and are predictive of engraftment in in

vivo assays and patients. A particularly preferred CAFC assay is described in Young, et al., Blood 88:1619 (1996). Flow cytometry can be used to subset hematopoietic cells from various tissue sources by the surface antigens they express. A combination of these assays may be used to test for hematopoietic cells or stem cells.

5 In one preferred embodiment the invention concerns a method of identifying genetically modified mammalian cells comprising introducing a nucleic acid sequence encoding a mutated protein-tyrosine kinase receptor (PTKR) as a selective marker and operatively linked to an expression control sequence into a target cell to form a genetically modified cell; allowing expression of the mutated PTKR in the genetically modified cell; and identifying said  
10 genetically modified cell expressing the mutated PTKR. In a preferred embodiment the selective marker is a mutated epidermal growth factor receptor (EGFR) family member, particularly preferred is a mutated EGFR1 and more specifically the sequences designated EGFR1-I and EGFR1-II. A polynucleotide or nucleic acid sequence is said to "encode" a polypeptide if, in its native state or when manipulated by methods well known to those of skill in the art it can be  
15 transcribed and/or translated to reproduce a polypeptide or fragment thereof. A construct or vector including a mutated PTKR of the invention may be introduced into the target population by any means of genetic transfer or modification known in the art.

The term "genetic modification" refers to any addition, deletion or disruption to a cells normal nucleotides and the methods of genetic modification are intended to encompass any  
20 genetic modification method of introducing nucleic acid sequences encoding the selective markers according to the invention and including but not limited to heterologous or foreign genes into mammalian target cells. These techniques are generally known. The term introducing is broadly used herein, and includes for example inserting. The term genetic modification includes but is not limited to transduction (viral mediated transfer of host DNA from a host or  
25 donor to a recipient, either in vivo or ex vivo), transfection (transformation of cells with isolated viral DNA genomes), liposome mediated transfer, electroporation, calcium phosphate transfection and others. Methods of transduction include direct co-culture of cells with producer cells (Bregni, et al., Blood 80:1418-1422 (1992)); culturing with viral supernatant alone, with or without appropriate growth factors and polycations (Xu, et al., Exp. Hemat. 22:223-230 (1994),  
30 and spinoculation.



In a preferred embodiment the mutated PTKRs of the invention and particularly the mutated EGFR markers are introduced into target cells by transduction with a retroviral vector as previously described. The host cell range that may be infected is determined by the viral envelope protein. The recombinant virus can be used to infect virtually any other cell type  
 5 recognized by the env protein provided by the packaging cell, resulting in the integration of the viral genome in the transduced cell and the stable incorporation of the foreign gene product. In general, murine ecotropic env of MoMLV allows infection of rodents cells, whereas amphotropic env allows infection of rodent, avian and primate cells including human cells. Amphotropic packaging of cell lines for use with retroviral systems are known in the art and are  
 10 commercially available. These include but are not limited to, PA12, PA317,  $\psi$ CRIP, and FLYA13. (See Miller, et al., Mol. Cell Biol. 5:431-437 (1985); Mill, et al., Mol. Cell Biol. 6:2895-2902 (1986); and Danos, et al. (1988) Proc. Natl. Acad. Sci. USA 85:6460-6464. Recently, the G-glycoprotein from vesicular stomatitis virus (VSV-G) has been substituted for the MoMLV env protein. (See Burns, et al., Proc. Natl. Acad. Sci. USA 90:8033-8037 (1993);  
 15 and WO92/14829). Xenotropic vector systems also exist which allow infection of human cells.

Once the target cells are genetically transformed with a nucleotide sequence including the mutated PTKR as a marker, and optionally a heterologous gene, the modified cells expressing the mutated PTKR selectable marker may be identified by numerous techniques. The term "identify or "identification" used herein in reference to genetically modified cells, unless  
 20 otherwise indicated means to mark, to purify, to enrich, to select, to isolate, or to separate. Identification can be by a single or multiple steps.

In a preferred embodiment, the genetically modified cells expressing the selective marker according to the invention are identified by an antibody that specifically recognizes and binds to the selectable marker, and particularly an antibody that recognizes a mutated EGFR.  
 25 This kind of antibody has been described by O'Rourke, et al., Oncogene, 16:1197-1207 (1998). A secondary antibody may then be used to further identify or select the antibody coated cells, if the secondary antibody is coupled to either a fluorophore or immuno-magnetic beads. The marker gene expressing cells may then be selected by flow cytometry or a by using a magnet to select bead-coated cells.

30 The techniques used for the identification of genetically modified cells expressing a selective marker of the invention include those described above and other well known

techniques including but are not limited to immunoselection; nucleotide detection by northern blots wherein RNA bound to a solid support is analyzed for binding to a liquid phase DNA; nucleotide detection by southern blots, wherein genomic DNA bound to a solid support is analyzed for binding to a liquid phase DNA; PCR amplification of genomic DNA; protein  
 5 detection by western blots whereby protein bound to a solid support is analyzed for binding to a liquid phase antibody; reverse transcription of mRNA and amplification with PCR; and FISH wherein chromosomes are analyzed by Fluorescence in situ hybridization with a liquid phase DNA (Lawrence, et al., Science, 249(4971):928 -932 (1990)).

The above listed methods are not described in detail herein; they are well known by  
 10 those of ordinary skill in the art. In brief, antibodies may be obtained by methods well known and reference is made to Harlow, et al., "Antibodies: A Laboratory Manual: (1988), Biosupplynet Source Book, (1999) Cold Springs Harbor Laboratory. Either polyclonal antibodies that are reactive to the antigen of interest may be used or monoclonal antibody producing cell clones may be generated. According to the invention the antibody must recognize  
 15 the extracellular domain of the mutated PTKR selective marker. More particularly if parts of the extracellular domain are modified, by for example deletion, the antibody should recognized an epitope of the remaining amino acid sequence of the mutated PTKR.

Monoclonal antibodies to EGFR are available commercially. Some sources include Calbiochem (CA), Pharmingen (CA), Becton Dickinson (CA), and the American Type Culture  
 20 Collection (ATCC) (Virginia). The antibody may be identified and assayed *in vitro* by a range of methods including gel diffusion, immunoassay, immunoelectrophoresis and immunofluorescence. Once genetically modified cells are labeled they can be incubated with an antibody against the mutated receptor.

The genetically modified cells may be physically separated by the use of antibodies,  
 25 flow cytometry including fluorescence activated cell sorting (FACS), and bead selection. (U.S. Pat. No. 5,011,912).

In FACS an antibody that recognizes the mutated PTKR may be used to identify the cells that express the PTKR. This primary antibody can be conjugated to a fluorophore, such as fluorescein isothiocyanate (FITC), phycoerythrin (PE), cy-chrome (CyC), allophycocyanine  
 30 (APC), tricolor (TC) or Texas Red (TX). If the primary antibody is not conjugated to a fluorophore a secondary antibody that is conjugated to a fluorophore may be introduced into the

cell sample containing cells which have expressed the mutated PTKR and which are recognized by the primary antibody. The primary antibody is attached to the mutant PTKR. The secondary antibody binds to the primary antibody and cells having the secondary antibody can be induced to fluoresce. Separation may be achieved by the fluorescence activated cell sorter.

5           The genetically modified target cells expressing a mutated PTKR, and optionally a second nucleic acid sequence encoding a protein of interest may be expanded, either prior to or after identification and selection, by culturing the cells for days or weeks in appropriate culture media, with or without supplements by means well known in the art. (Freshney *supra*, Celis *supra*, and Coligan et al., *supra*).

10           The genetically modified marked cells obtained according to the methods of the invention may further be used in an autologous or allogeneic setting wherein the genetically modified target cells, preferably hematopoietic cells are expanded and then used in gene therapy for example in bone marrow transplantation, graft facilitation, or immune reconstitution. The expanded cells expressing the mutated cell surface marker may be infused into a subject.

15           Samples may be taken and then tested for the selectable marker by FACS analysis, PCR, or FISH, as referenced above, to determine the persistence of the marked cells and further to assess efficiency of transduction.

            The invention generally described above will be more readily understood by reference to the following examples, which are hereby included merely for the purpose of illustration of

20           certain embodiments of the present invention and are not intended to limit the invention in any way.

### **Example 1:**

#### 25    **A. Isolation of human EGFR cDNAs:**

EGFR cDNA is isolated by PCR from cDNA that has been generated from colorectal adenocarcinoma: SW480 cell line (Marathon cDNA, Clontech (CA); Leibowitz, et al., Cancer Research 36:4562-4569 (1976). The following primers are used in the described methods and were obtained from Life Technologies (MD):

30

5' Primers:

EGFR1 CTA GGC TAG CAT GCG ACC CTC CGG GAC GGC C SEQ ID NO. 3

EGFR2 CTC TGC CCG GCG AGT CGG GCT GAC AGC TAT GAG ATG GAG GAA SEQ ID  
NO. 4

5 3' Primers:

EGFR3 TTC CTC CAT CTC ATA GCT GTC AGC CCG ACT CGC CGG GCA GAG SEQ ID  
NO. 5

EGFR2220R GGA TAT CCT ACG TGC GCT TCC GAA CGA TGT G SEQ ID NO. 6

10 B. Generation of EGFR1-I:

The primers EGFR1 and EGFR2220R are used to generate an intracellular deletion mutant of EGFR (Figures 2 and 3) from the SW480 adenocarcinoma cell line cDNA. Amino acid residues 679 – 1210 of the sequence in Figure 2 (SEQ ID NO. 2) are deleted, and this mutant EGFR is designated EGFR1-I.

15 The 5' primer EGFR1 covers the start codon of the EGFR. The 3'-primer EGFR2220R contains a stop codon in place of amino acid 679 of the SEQ ID NO. 2 (Figure 2). Using primers EGFR1 and EGFR2220R results in the amplification of the EGFR sequence that has a deletion of amino acids 679 to 1210.

The following PCR reaction is performed: Marathon cDNA (~ 0.5 ng) is mixed with  
20 Advantage cDNA buffer (10 mM Tris-HCl (pH=7.5 at 42°C), 50 mM KCl, 2.5 mM MgCl<sub>2</sub>, 0.001% Gelatin, 2.5 µmol dATP, 2.5 µmol dCTP, 2.5 µmol dGTP, 2.5 µmol dTTP), 0.01 OD<sub>260</sub> primer EGFR1, 0.01 OD<sub>260</sub> primer EGFR2220R, 1 µl Advantage cDNA polymerase (Life Technologies; MD), 5U Pfu Turbo polymerase (from *Pyrococcus furiosus*) (Promega; WI) and water in a final volume of 50 µl. PCR is performed as follows: Cycle 1: 95°C for 5 min, Cycle  
25 2-15: 95°C for 1 min, 60°C for 1 min, 68°C for 4 min, and Cycle 16: 68°C for 10 min.

The reaction is then cooled to 4°C in the PCR machine and subsequently the amplified cDNA is ethanol precipitated with 0.3 M sodium acetate. The pellet is washed once with 70% ethanol, dried and resuspended in 50 µl H<sub>2</sub>O.

1 µl of the above PCR reaction is then reamplified, and the reaction mix contains for the  
30 second round of amplification in addition to 1 µl of the above PCR reaction: Pfu buffer (20mM Tris-HCl (pH8.8), 2 mM MgSO<sub>4</sub>, 10 mM KCl, 10 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.1 % Triton X-100, 0.1

mg/ml BSA), 2.5  $\mu$ mol of each dNTP (dATP, dTTP, dCTP, dGTP), 0.01 OD<sub>260</sub> primer EGFR1, 0.01 OD<sub>260</sub> primer EGFR2220R, 5 U Pfu Turbo Polymerase and water in a final volume of 50  $\mu$ l. PCR is performed as follows: Cycle 17: 95°C for 5 min, Cycle 18–48: 95°C for 1 min, 60°C for 1 min, 72°C for 4 min, and Cycle 49: 72°C for 10 min.

5 The reaction is cooled to 4°C in the PCR machine and the amplified cDNA is ethanol precipitated with 0.3 M sodium acetate. The pellet is washed once with 70% ethanol, dried and resuspended in 20  $\mu$ l H<sub>2</sub>O. The PCR reaction is loaded on a 1xTAE gel. A band with the size of approximately 2200 bp is isolated from the gel and cloned into the SrfI restriction site of pPCR-Script Amp vector (Stratagene; CA) according to the manufacturer's protocol. The  
10 resulting vector is called pPCR-Script EGFR1-I. The correctness of the subcloned PCR product is confirmed by restriction analysis and sequencing according to well known methods.

### 15 C. Generation of EGFR1-II:

EGFR1-II contains the same deletion in the intracellular domain as EGFR1-I (a stop codon in position 679). In addition, amino acids 25 to 312 are deleted in the extracellular domain of EGFR depicted in Figure 2 (SEQ ID NO. 2). The sequence of the signal peptide, amino acid 1 to 24, is fused to amino acid 313 of the EGFR extracellular domain. In order to  
20 delete the EGFR extracellular domain, a protocol is used that is described in further detail below. (Also see White, ed. Methods in Molecular Biology, Vol 15. PCR Protocols, Chapter 25 (1993)). This method is known in the art as “gene splicing by overlap extension (gene SOEing)”. (Figure 4)

Primer pairs EGFR1/EGFR3 and EGFR2/EGFR2220R are used to amplify 2 PCR  
25 products (a) and (b). Primer EGFR1 encodes amino acid 1 to 7, and Primer EGFR3 encodes amino acid 18 to 24 fused to the nucleotide sequence of amino acid 313 to 319. Using the EGFR1 and EGFR3 primers a PCR product is generated that encodes amino acid 1 to 24 which is fused to amino acid 313 to 319.

Primer EGFR2 encodes the same amino acid acids as Primer EGFR3 only in the reverse  
30 orientation. Primer EGFR2220R encodes a stop codon in place of amino acid 679. As

described in section B above, this results in the deletion of the intracellular domain after amino acid 678.

Using these two primer pairs, two PCR products (a) and (b), are obtained each having a 42 nucleotide of overlapping sequence encoding amino acid 18 to 24 and amino acids 313 to 319 (Figure 4 – PCR reaction 1).

In the second PCR reaction, primers EGFR1 and EGFR2220R are used with the intermediate PCR products (a) and (b). When these intermediate PCR products are mixed, denatured and reannealed, one of the strands of the two PCR products can overlap at their 3' ends, and act as primers on one another to make the mutant product. Subsequently, the mutant PCR product (c) is amplified with the primers EGFR1 and EGFR2220R. The resulting PCR product encodes EGFR1-II amino acid 1 to 24 of the wt EGFR fused to amino acids 313 to 678 of the wt EGFR with a stop codon in position 679 (Figure 4).

The PCR reactions are as follows: The first PCR reaction in: a.) 0.01 OD<sub>260</sub> of EGFR1 and 0.01 OD<sub>260</sub> of EGFR3 and in b.) 0.01 OD<sub>260</sub> EGFR2 and 0.01 OD<sub>260</sub> EGFR2220R are mixed with ~0.4 ng Marathon cDNA, 1x Advantage cDNA Polymerase buffer (as described in B), 1 µl Advantage cDNA polymerase, 2.5 µmol of each dNTP, 5U Pfu Polymerase (Promega) and H<sub>2</sub>O in a final volume of 50 µl. The PCR is performed as follows: Cycle 1: 94°C for 5 min, Cycle 2-16: 94°C for 0.5 min, 60°C for 1 min, 68°C for 7 min, and Cycle 17: 68°C for 10 min.

The reactions are cooled to 4°C in the PCR machine and the amplified PCR products are ethanol precipitated with 0.3 M sodium acetate. The pellets are washed once with 70% ethanol, dried and resuspended in 50 µl H<sub>2</sub>O.

5 µl of each PCR reaction (a) and (b) are then further amplified. The reaction mix contains for the second round of amplification, in addition to each PCR reaction (a) 0.01 OD<sub>260</sub> of primers EGFR1 and EGFR3 or (b) 0.01 OD<sub>260</sub> of primers EGFR2 and EGFR2220R: 1 x Pfu buffer, 2.5 µmol of each dNTP, and 5 U Pfu Polymerase (Promega; WI) and H<sub>2</sub>O in a final volume of 50 µl. The PCR is performed as follows: Cycle 18: 94°C for 5 min, Cycle 19–49: 94°C for 0.5 min, 60°C for 1 min, 72°C for 6 min and Cycle 50: 72°C for 10 min.

The reactions are then cooled to 4°C in the PCR machine and the amplified PCR products are phenol/chloroform extracted and ethanol precipitated with 0.3 M sodium acetate. The pellet is washed once with 70% ethanol, dried and resuspended in 20 µl H<sub>2</sub>O.

In a second PCR reaction, PCR product (a) and (b) are mixed in equimolar amounts and reamplified with Pfu Turbo Polymerase (Stratagene) using the primer pair EGFR1 and EGFR2220R. The PCR mix is the same as described for cycles 18-50. The PCR is performed as follows: Cycle 1: 95°C for 5 min, Cycle 2 – 32: 95°C for 1 min, 60°C for 1 min, 72°C for 4 min, and Cycle 33: 72°C for 10 min. The PCR is then loaded on a gel (as described above) and a band of ~1200 bp is isolated. This PCR product is cloned into the SrfI site pPCR-Script (Stratagene; CA) according to the manufacturer's protocol. The resulting vector is called pPCR-Script EGFR1-II. The correctness of the subcloned PCR product (c) is confirmed by restriction analysis and sequencing according to well known methods.

#### D. Generation of retroviral vectors containing mutated EGFRs and viral supernatants:

EGFR1-I and EGFR1-II sequences are excised from pPCR-Script EGFR-I and pPCR-Script EGFR-II using the restriction sites NotI and XhoI and are cloned into the multiple cloning site of the Murine Leukemia Virus (MoMLV) based retroviral vector pG1a (GTI, Maryland) – (pG1a-mutated sequence – IRES- nerve growth factor receptor (NGFR)), which has been cut with NotI and XhoI. The retroviral vectors are designated pG1aEGFR1-I and pG1aEGFR1-II.

The constructs are cotransfected into human embryonic kidney cells 293T (293T cells) (Gary Nolan, Stanford) with an envelope construct pCiGL that permits expression of the Vesicular Stomatitis Virus G-Protein (VSV-G envelope) under the control of the cytomegalovirus (CMV). Also cotransfected into 293T cells is the packaging construct pCiGP (encoding MoMLV gag-pol under the control of the CMV promoter) using the CaCl<sub>2</sub> technique (Clontech; CA). (WO 97/21825 and Rigg, et al., Virology 218:290-295 (1996))

Viral supernatants are collected 24, 48, and 72 hours after transfection, centrifuged at 1200 rpm in a Beckman GS-6KR centrifuge to remove particulate material, and either used immediately to transduce cells or frozen in a dry ice/methanol bath. The viral supernatants are used to transduce the packaging cell line ProPak-A-6 (PPA-6) (Systemix, Inc.). The PPA-6 cell line is a derivative of 293T cells expressing the MLV amphotropic envelope and MLV gag/pol under the control of the CMV promoter (Rigg et al. supra). The PPA-6 cells are sorted by immuno-magnetic bead selection. Alternatively, if the cells only expressed low amounts of mutated EGFR, the cells are sorted by FACS sorting as described in detail below. Supernatants from PPA-6 cells are collected on day 2, 3 and 4 after transduction and treated as described for

293T cells. The generated supernatants of PPA-6 cells contain recombinant viral particles that have the amphotropic envelope and are used to transduce human primary cells and cell lines.

#### E. Tissue culture and cell lines:

5           The following cell lines and primary cells are used: (a) human T cell line, CEMSS (Frederico, et al., J.Biol. Regul. Homeost Agents, 7: 41-49 (1993)), (b) human embryonic kidney cells 293T (293T) (Pear, et al., Proc. Natl. Acad. Sci. USA 90:8392 - 8396 (1993)), and (c) PPA-6 (Rigg, et al., supra). Human primary T cells are obtained by isolating the mononuclear cell fraction from human blood using Ficoll-gradient centrifugation (Noble and  
10   Cutts, Can. Vet. J. 8: 110 – 111 (1967) and Boyle and Chow, Transfusion 9:151 – 155 (1969). Human blood (buffycoats; Pietersz, et al., Vox Sang 49: 81 – 85 (1985) and Pietersz, et al., Blut. 54:201 – 206 (1987) obtained from the Stanford Bloodbank are diluted 1:1 with phosphate buffered saline (PBS). In 50 ml tubes, 15 ml blood are overlaid onto 21 ml of Ficoll (Isoprep; Robbins Scientific, California). The gradient is spun for 30 minutes at 1700 rpm in a Beckman  
15   GS-6KR at room temperature and stopped without using the break. Peripheral blood mononuclear cells (PBMCs) are collected from the interface of the Lymphoprep gradient. The resulting PBMCs are further purified on a second Ficoll gradient to remove remaining blood cells. The PBMCs are incubated for 1 hour in a tissue culture flask in tissue culture media for PBMCs (see below) at 37°C in 5% CO<sub>2</sub> allowing adherent cells such as macrophages to attach  
20   to the tissue culture flask. Non-adherent cells (T/B/NK cells) are used to purify CD4<sup>+</sup> T cells. Cells (2 x 10<sup>8</sup>) are incubated with 300 µl anti-CD4 antibody for 1 hour at 4°C. The cells are washed 3 times with PBS and incubated with 1 x 10<sup>8</sup> anti-mouse IgG coupled immuno-magnetic beads (Dyna, Oslo) for 1 hour at 4°C. The CD4<sup>+</sup> cells that bound the CD4-antibodies and the secondary antibody coupled magnetic beads are positively selected by incubating the cells for  
25   10 minutes on a Dynal magnet. After removing the unbound cells, the remaining cells are taken off the magnet and put into culture (see below). Usually the beads remain for up to 10 days on the cells.

CD34<sup>+</sup> cells are isolated from G-CSF mobilized peripheral blood (MPB) using Isolex 300SA or 300I (Baxter, IL) (Systemix, CA). The cells are approximately 80 – 90% pure.

30           Cells are cultured in a Steri-Cult 200 incubator (Forma-Scientific) at 5% CO<sub>2</sub>. Media (DMEM, Iscove's medium RPMI), PBS, and sodium pyruvate are obtained from JRH



Biosciences (California), FBS from Hyclone (Utah), L-glutamine, Trypsin and MEM vitamins from Life Technologies (Maryland), ITS (insulin/transferrin/sodium selenite), PHA (phythemagglutinin), Interleukin-2 (IL-2) from Sigma (Missouri).

The cells are cultured in DMEM, 10% FBS, 1% sodium pyruvate, 1% L-glutamine (293 T, PPA-6); RPMI, 10% FBS, 1% sodium pyruvate, 1% L-glutamine (CEMSS), Iscove's medium 10% FBS, 1% L-glutamine, 1% ITS (insulin/transferrin/sodium selenite, 0.5 mg/ml stock), 1% MEM vitamins (human PBMCs). CD4<sup>+</sup> T cells are stimulated every 10-12 days with 1-2 µg/ml PHA, irradiated feeder cells and 40 U/ml IL-2.

Irradiated feeder cells are generated as follows: PBMCs are isolated as described above and irradiated with 3500 rads. The cells are mixed at a 10:1 ratio with the B cell lymphoma JY (Barbosa, et al., Proc. Natl. Acad. Sci. 81:7549-7553 (1984)) that had been irradiated with 6000 rads.

To obtain T-cells in a resting state, 4-5 days after stimulation T cells are transferred into fresh medium that contained 20 U/ml IL-2. CD34<sup>+</sup> cells are grown in X vivo 15, 50 ng/ml thrombopoietin (TPO; R & D Systems, MN), 100 ng/ml flt-3 ligand (FL), 100 ng/ml steel factor (SF). (SyStemix, Inc.)

In order to passage adherent cells, the cells are washed once with PBS, then trypsinized for 5 min and subsequently split into new tissue culture flasks (VWR, New Jersey).

#### 20 F. Transduction of PPA-6, human T cell lines, primary T cells and CD34<sup>+</sup> cells:

10<sup>6</sup> cells/ml from step (E) are transduced with 1-3 ml of viral supernatant, that has been either generated from 293 T cells or PPA-6 cells, by spinoculation with 8 µg/ml protamine sulfate (Sigma, Missouri). Spinoculation is done at 37°C for 3 hrs at 3000 rpm (CD34<sup>+</sup> cells, PPA-6) or at 2750 rpm (human T cells, CEMSS). PPA-6 cells are transduced in 6 well plates, non adherent cells in 6 ml tubes (VWR). Human T cells and CD34<sup>+</sup> cells are activated for 2 25 days prior to the transduction procedure with PHA/IL-2 or TPO/FL/SF, respectively.

#### G. FACS analysis, FACSorting and immuno-magnetic bead selection of cells expressing EGFR1-I and EGFR1-II:

30 The following antibodies and reagents are used for staining. CD4-FITC (Caltag; CA), anti-CD34-APC (Becton Dickinson; NJ), Thy-1-PE (Becton Dickinson), propidium iodide (PI)

(Sigma), unconjugated anti-EGFR antibody (GR01) (Calbiochem; CA), goat anti-mouse IgG-PE/FITC (Caltag), goat anti-mouse IgG<sub>2a</sub>-PE/FITC (Caltag), and anti-mouse IgG coupled magnetic beads (Dyna; Oslo). The anti-bodies and immuno-magnetic beads were used according to manufacturer's protocol.  $1 \times 10^6$  cells are stained in 50  $\mu$ l of PBS/2% FBS for 20 to 60 minutes at 4°C. When a secondary antibody is used the cells are washed once with 2 ml of PBS/2%FBS, then incubated again in 50  $\mu$ l PBS/2%FBS and the secondary antibody is added. Before FACS analysis, the cells are again washed once with PBS/2%FBS, centrifuged and then resuspended in 500  $\mu$ l PBS/2%FCS containing 1  $\mu$ g/ml PI. FACS analysis is performed on a FACScan (Becton-Dickinson Immunocytometry Group; CA) according to manufacturer's instructions.

Cells that express the marker genes, EGFR1-I and EGFR1-II are sorted as follows:  $2 \times 10^7$ /ml PBS/2%FBS cells are stained with an anti-EGFR antibody (at least with 1  $\mu$ l antibody/ $10^6$  cells; GR01). The cells are washed twice with PBS/2%FBS and subsequently incubated with a fluorophore (FITC, PE) conjugated rat anti-mouse IgG antibody (5  $\mu$ l/ $10^6$  cells) that recognizes the anti-EGFR antibody. The cells are also stained with PI to distinguish between dead and live cells.  $1 \times 10^6$  cells/ml are sorted for EGFR-positive and PI-negative cells. This is done on a FACStar Plus (Becton-Dickinson Immunocytometry Group) according to manufacturer's protocol.

To isolate cells by bead selection, the cells are stained with an anti-EGFR antibody. The  $10^7$  cells/ml are incubated with 10  $\mu$ l/ml anti-EGFR antibody (GR01) in PBS/2%FBS for 1 hr on ice with occasional shaking. The cells are washed 3 times with PBS/2%FCS and then anti-IgG antibody coupled magnetic beads (Dyna; Oslo) that can recognize the anti-EGFR antibody are added (~ 5 beads per positive cell). Again the cells are incubated for 1 hr on ice. The cells that express EGFR are selected by positive selection with a Dynal magnet (Dyna; Oslo) for 10 min. The unbound cells are removed and the EGFR expressing cells are put into culture (see section E.).

Figure 5 illustrates the expression of EGFR1-II CD34<sup>+</sup> cells after transduction of these cells with supernatants that are generated from PPA-6 cells. Primary human T cells are transduced with PPA-6 supernatants and are selected by immuno-magnetic beadselection then sorted using a FACStar Plus (data not shown). The cells are enriched from 16 % to 92 %.

**Example 2:****A. Isolation of human MuSK-R cDNAs:**

MuSK-R is isolated by PCR from fetal skeletal muscle cDNA (Marathon cDNA, Invitrogen)  
5 using primers flanking the 5' and 3' of the MuSK-R cDNA.

The following primers obtained from Operon Technologies, Inc. are used to amplify MuSK-R cDNA:

	MuSK21FN: CGT CCT GCG TGA GCC TGG ATT AAT C	SEQ ID NO: 9
	MuSK34FN: GCC TGG ATT AAT CAT GAG AGA GCT C	SEQ ID NO: 10
10	MuSK2666RN: CGA GGC CTG TCT TCA ACC TTA GAC	
	ACT CAC AGT TCC CTC TGC	SEQ ID NO: 11

The 5' primer MuSK21FN covers 25 nucleotide (nt) before the start codon, the second 5' primer MuSK34FN covers the start codon (aa 1) of MuSK-R and surrounding sequence. The 3'-primer MuSK2666RN covers the stop codon of MuSK-R and surrounding sequence. Using  
15 primers MuSK21FN, MuSK34FN and MuSK2666RN results in the amplification of a DNA fragment of ~ 2600 bp that encodes a MuSK-R wt protein.

The following PCR reaction is performed: Marathon cDNA (~ 2 ng) is mixed with Advantage cDNA buffer (10 mM Tris-HCl (pH=7.5 at 42°C), 50 mM KCl, 2.5 mM MgCl<sub>2</sub>, 0.001% Gelatin), 2.5 µmol dATP, 2.5 µmol dCTP, 2.5 µmol dGTP, 2.5 µmol TTP), 1 µg primer  
20 MuSK21FN, 1 µg primer MuSK2666RN, 1 µl Advantage cDNA polymerase, and water in a final volume of 50 µl. The PCR is performed as follows: Cycle 1: 94°C for 5 min, Cycle 2-11: 94°C for 0.5 min, 63°C for 1 min, 68°C for 6 min, and Cycle 12: 68°C for 10 min.

The reaction is cooled to 4°C in the PCR machine, and the amplified cDNA is ethanol precipitated with 0.3 M sodium acetate. The pellet is washed once with 70% ethanol, dried and  
25 resuspended in 100 µl H<sub>2</sub>O.

10 µl of the above PCR reaction is then reamplified. The reaction mix contains for the second round of amplification step in addition to 10 µl of the above PCR reaction: Pfu buffer (20mM Tris-HCl (pH8.8), 2 mM MgSO<sub>4</sub>, 10 mM KCl, 10 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.1 % Triton X-100, 0.1 mg/ml BSA), 2.5 µmol of each dNTP (dATP, dCTP, dGTP, dTTP), 1 µg primer  
30 MuSK34FN, 1 µg primer MuSK2666RN, 5 U Pfu Turbo Polymerase (from *Pyrococcus furiosus*) and water in a final volume of 50 µl. The PCR is performed as follows: Cycle 13:

94°C for 5 min, Cycle 14–43: 94°C for 0.5 min, 62°C for 1 min, 72°C 6 min, and Cycle 44: 72°C for 10 min.

The reaction is cooled to 4°C in the PCR machine and the amplified cDNA is ethanol precipitated with 0.3 M sodium acetate. The pellet is washed once with 70% ethanol, dried and resuspended in 20 µl H<sub>2</sub>O. The PCR reaction is loaded on a 1xTAE gel. A band with the size of ~ 2600 bp is isolated from the gel and cloned into the SrfI restriction site of pPCR-Script Amp vector (Stratagene, CA) according to the manufacturer's protocol. The resulting vector is called pPCR-Script MuSK-R-wt. The correctness of the subcloned PCR product is confirmed by restriction analysis and sequencing by methods well known in the art. (The nucleotide sequence is illustrated in SEQ ID NO.7)

*B. Generation of mutations in the intracellular domain of MuSK-R by PCR:*

The primers MuSK1380F, MuSK1657R, and 1747R are used to generate intracellular deletion mutants of MuSK-R from the plasmid pPCRScriptMuSK-R. The primer sequences are as follows wherein p means phosphorylated:

Primer 1380F: 5' pCG GCC TGT GCC AGA CTG CCA CAT CTA G (SEQ ID NO: 12);

Primer 1657R: 5' pCG TCT AGG TGA GGG TTA CTG CTG  
CTG ATT CTC (SEQ ID NO:13)

Primer 1747R: 5' pGG TTA ACC CTA TTC AAT GTT ATT  
CCT TGA ATA CTC CAG (SEQ ID NO: 14).

Using primer pair MuSK1380F and 1657R results in the deletion of amino acid residues 538 – 879 of MuSK-R, using primer pair MuSK1380F and 1747R results in the deletion of amino acid residues 577 - 879. The two mutant forms of MuSK-R are designated MuSK-RA538-879 (MuSK-RI) and MuSK-RA577-879 (MuSK-RII). In both MuSK-RI and MuSK-RII most of the intracellular domain of MuSK-R as shown in Figure 6 is deleted. While not meant to limit the invention in any manner, it is believed that both truncations result in a deletion of the kinase domain and most of the substrate binding motifs of the wt MuSK-R illustrated in Figure 6.

The 5' primer MuSK1380F covers the nucleotide sequence 1333-1410 of the MuSK-R. The 3'-primers MuSK1657R and 1747R contain stop codons in place of amino acid 538 and 577 of MuSK-R. Using primer MuSK1380F with MuSK1657R or MuSK1747R results in the amplification of MuSK-R nucleotide sequence 1333 to 1614 that has a stop codon in the position of amino acid 538 or nucleotide sequence 1333-1728 that has a stop codon in the position of amino acid 577, respectively. The PCR reaction includes ~10 ng hMuSK-R wt DNA, 1 x Pfu buffer, 1 µg of primer MuSK1380F and either 1 µg primer MuSK1657R or MuSK1747R, 2.5 µmol of each dNTP, 5 U Pfu polymerase and H<sub>2</sub>O in a final volume of 50 µl. The PCR reaction is performed as follows: Cycle 1: 95°C for 5 min, Cycle 2-31: 95°C 0.5 min, 60°C for 1 min, 72°C for 4 min, Cycle 32: 72°C for 10 min. The PCR reaction is cooled to 4°C in the PCR machine and then loaded on a 1 x TAE gel.

The two PCR products MuSK-RI (nt 1380-1614) and MuSK-RII (nt 1380 – 1728) are cloned into the SrfI site of pPCR-ScriptAmp (Stratagene) according to the manufacturer's protocol. As the 5' coding sequence from MuSK (nt 1 – 1379) is missing in these constructs, this sequence is excised from the plasmid pPCR-Script MuSK-wt using restriction sites NaeI and AatII. The two pPCR-Script vectors containing the modified MuSK sequence nt 1- 1614 and 1 – 1726 are called pPCR-Script-MuSK-RI and pPCR-Script MuSK-RII, respectively. The correctness of the vectors are confirmed by restriction analysis and sequencing by methods well known in the art.

#### C. Generation of retroviral vectors containing mutated MuSK-Rs and viral supernatants:

Wild-type and mutant MuSK-R are excised from pPCRScriptMuSK-Rwt, pPCRScriptMuSK-RI and pPCRScriptMuSK-RII using the NotI and XhoI site and are cloned into the multiple cloning site of the Moloney Murine Leukemia Virus (MoMLV) based retroviral vector pG1a (GTI, Maryland) which is cut with NotI and XhoI. The retroviral vectors are designated pG1aMuSK-R, pG1aMuSK-RI and pG1aMuSK-RII. The constructs pG1aMuSK-R, pG1aMuSK-RI and pG1aMuSK-RII were used to generate viral supernatants as described in Example 1.

D. Tissue culture and cell lines: Experiments were carried out the same way as described in Example 1.

E. Transduction of PPA-6 and human T cell line: Experiments were carried out the same way as described in Example 1

5 F. FACS analysis and immuno-magnetic bead selection of cells that express MUSK-RI and MUSK-RII:

In addition to the antibodies that are listed in Example 1, anti-MuSK-R polyclonal serum, and anti-MuSK-R hybridoma supernatant (see section G) were used for flow cytometry and immuno-magnetic bead selection. .

10 In order to isolate cells by immuno-magnetic bead selection, the cells are stained with an anti-MUSK-R antibody. For this purpose the  $10^7$  cells/ml are incubated with 1-3 ml anti-MuSK-R hybridoma supernatant in PBS/2%FBS for 1 hr on ice with occasional shaking. The cells are washed 3 times with PBS/2%FCS and then anti-IgG antibody coupled magnetic beads, that can recognize anti-MuSK-R antibodies, are added (~ 5 beads per positive cell). The cells are  
15 incubated for 1 hr on ice. Cells that express MuSK-R are selected by positive selection with a Dynal magnet (Dynal, Oslo) for 10 min. The unbound cells are removed and the MuSK-R expressing cells are put into culture as described in section D.

G. Generation of a monoclonal antibody against the extracellular domain of MuSK-R:

20 To generate monoclonal antibodies against the extracellular domain (XC) of MuSK-R, MuSK-R XC is amplified by PCR and cloned into the expression construct pSecTag2b (Invitrogen). Cloning the XC domain of MuSK-R into the multiple cloning site (MCS) of the plasmid pSecTag2b allows for the expression of the XC under the control of the CMV promoter. In addition, the plasmid contains the sequence of a myc and (His)<sub>6</sub>-tag after the multiple cloning  
25 site, which allows to fuse the protein of interest (MuSK-RXC) to the myc and (His)<sub>6</sub>-tag. The signal peptide of MuSK-R is replaced by the Igκ leader. (Figure 3). The extracellular domain of MuSK-R without the signal peptide is amplified by PCR using the following primers wherein p means phosphorylated:

MuSK 116FPC: 5' pCT TCC AAA AGC TCC TGT CAT CAC C SEQ ID NO: 15

30 MuSK 1532RC: 5' pCC AGT CAT GGA GTA TGT AGG TGA GAC SEQ ID NO: 16

Primer MuSK116FPC starts with the sequence after the signal peptide (nt 69 – 93). Primer MuSK1532RC covers the sequence before the transmembrane domain starts and the first 2 amino acids of the transmembrane domain corresponding to nucleotide sequence 1462-1586 of SEQ I NO:7. For the PCR reaction ~10 ng hMuSK-R wt DNA are mixed with 1 x Pfu buffer, 1 µg of primer MuSK116FPC and 1 µg primer MuSK1532RC, 2.5 µmol of each dNTP, 5 U Pfu polymerase and H<sub>2</sub>O in a final volume of 50 µl. The PCR reaction is performed as follows: Cycle 1: 95°C for 5 min, Cycle 2-7: 96°C for 0.5 min, 60°C for 1 min, 72°C for 6 min, and Cycle 8 - 27: 95°C for 0.58 min, 58°C for 1 min, 72°C for 6 min; Cycle 28: 72°C for 10 min. The PCR reaction is cooled to 4°C in the PCR machine and then gel-purified. The PCR fragment is cloned into to the EcoRV restriction site of pSecTag2b. In this way MuSK-R XC is cloned in frame with the Igk leader at the N-terminus and the myc- and (His)<sub>6</sub>-tag at the C-terminus. The resulting plasmid is called pSecTag-hMuSK-R.

To express the MuSK-R XC the plasmid pSecTag-hMuSK-R is transfected into 293T cells by the CaCl<sub>2</sub> technique (as described in section C). 24 hrs after the transfection, the media is replaced with either fresh DMEM/10%FBS or serum-free X-Vivo 15. Supernatants of the transfected cells are collected after 48 and 72 hrs. A total of 400 ml supernatants are collected and are frozen at -80°C until the supernatants are purified.

The MuSK-R XC is purified from tissue culture supernatants by immobilized metal affinity chromatography. The metal ion is 0.1 M NiCl<sub>2</sub>. The column is a 1 or 5 ml Pharmacia metal HiTrap chelating sepharose column. The equilibration buffer (Buffer A) consisted of 20 mM Na<sub>2</sub>HPO<sub>4</sub> pH 7.4, 1M guanidine hydrochloride, 1M NaCl, filtered through 0.2 µM cellulose acetate filter. The elution buffer (Buffer B) is 20 mM Na<sub>2</sub>HPO<sub>4</sub>, pH 7.4, 1 M guanidine hydrochloride, 1M NaCl, 0.5 M imidazole, filtered through 0.2 µM cellulose acetate filter. The Pharmacia FPLC chromatography system is used to run columns, with FPLC director program software and a Pharmacia P50 pump. The purification is performed at 4°C.

The pump is primed with buffer A before the load is started. Before the column the column is attached, the load is pumped through until the pink color of the tissue culture media is seen at the connection so that the column is not washed with non-equilibration conditions.

The tissue culture supernatants are adjusted to contain 0.85 M NaCl, 1M guanidinium chloride and 40 mM imidazole and the pH is adjusted to 7.4. The column is equilibrated with 8% buffer B. The sample is loaded and the column then washed in above conditions for seven

column volumes. MuSK-R is eluted at 30% Buffer B (150 mM imidazole) over eight column volumes. Fractions are collected from start of the run.

Each fraction is tested in a Dot Blot Assay (see below). Selected positive fractions are tested in Western Blot assays and Elisa (see below). Positive fractions are pooled and dialyzed in 10,000 MWCO membrane (Pierce Snakeskin) against PBS. After dialysis, the optical density is determined at OD<sub>280</sub>. The samples are filtered through 0.2 µm filters and then concentrated in Centricon Centriprep 30 devices in a refrigerated Sorvall RT6000D according to manufacturer's protocol.

To test for positive samples in the dot blot assay, 10 µl of each fraction is pipetted on nitrocellulose. The membrane is dried, blocked with superbloc and then probed for MuSK-R protein and developed as described for the India Western (see below).

Western Blots are performed by methods well known in the art. The materials (sample buffer, running buffer and gels) are obtained from Novex.

For western blot analysis 25 to 35 µl/fraction are used. Guanidine containing fractions are precipitated in ice-cold ethanol and are stored on ice for 15 minutes or overnight at 4°C. The samples are centrifuged in refrigerated microcentrifuge (TOMY) at 14,000 rpm for 10 min. The supernatant is discarded, ice-cold acetone is added and centrifuged as before. The pellet is resuspended in SDS sample buffer (Novex) with 5% β-mercaptoethanol in a final volume of 50-70 µl. The samples are denatured at > 90°C for 5 minutes, briefly centrifuged, and 25 - 35 µl loaded on a 4-20% gradient gel. The gel is blotted onto 0.45 µm nitrocellulose for 1.4 hours at 100 volts, using the Biorad wet transfer blotting cassette with tris-glycine-methanol transfer buffer (25 mM trizma Base, 192 mM glycine, 20% methanol). After blotting the gel, the blot is blocked in Pierce TBS superbloc for 10 minutes with mild agitation. The blot is washed twice in TBST (50 mM Tris, pH 7.5, 150 mM NaCl, 0.05% Tween 20) for five min per wash on a rotating platform. The Pierce India™ His-HRP Probe is diluted to 1:5000 in TBST and the blot is incubated with the probe for 1 hour at room temperature and washed 3 times in TBST. After that, horseradish peroxidase reagent (Sigma Fast HRP Insoluble Substrate D4418) is added to the blot and the blot is developed. The blot is washed in three changes of water to stop development. Alternatively an mouse anti-c-myc antibody, (Santa Cruz Biotechnology; CA) is used to detect recombinant MuSK-R protein. This antibody is diluted in superbloc to 1 µg/ml final concentration. The blot is washed three times with TBST and then a goat anti-mouse IgG-



HRP antibody (Sigma) is added at 1:5000 dilution in superbloc. The blot is incubated for 1 hour at room temperature, with gentle agitation and developed as described above with Fast HRP insoluble substrate (Sigma). The recognized protein traveled at about 85 kD on the SDS PAGE, and it is considered to be 19 kD heavier due to glycosylation.

5        The recombinant MuSK-R protein is injected into 3 different Balb/c mice. For this purpose 25-50  $\mu$ g are mixed with 2.25 mg alhydrogel and 100  $\mu$ g MDP (muranyl dipeptide; Pierce) in a final volume of 200  $\mu$ l and injected 5 times every 14 days subcutaneously. After the 3<sup>RD</sup> and the 5<sup>TH</sup> injection serum of the 3 mice are tested for reactivity against MuSK-R by FACS analysis and Elisa. For FACS analysis the  $5 \times 10^5$  cells of cell lines CEMSS and CEMSSMuSK-R are used. Both the pre-sera and sera are diluted 1:100, 1:300, 1:900 and 1:2700. A rat anti-mouse IgG-PE antibody is used as a secondary reagent at a 1:20 dilution. For the Elisa, 96 well plates are coated with 50  $\mu$ l of 10  $\mu$ g/ml anti-mouse IgG<sub>Fc</sub> (Jackson; Maine). The plates are incubated with various dilutions of sera (1:100 to 1:218700), subsequently with MuSK-R protein and with Nickel activated horse radish peroxidase at a 1:1000 dilution (HRP, Pierce).

10        Nickel activated HRP is binding to the recombinant MuSK-R protein via the (His)<sub>6</sub> tag. To develop the Elisa, the plate is incubated with TMP peroxidase substrate (Zymed; CA). In both assays one mouse shows the highest reactivity against native and recombinant MuSK-R. This mouse is boosted with a 6<sup>th</sup> injection of 200  $\mu$ g MuSK-R protein in PBS. The injection is done subcutaneously and intravenously. 1 week later the spleen is removed, lymphocytes isolated

15        with lympholite M (Accurate Chemicals) and fused, using 50% polyethylene glycol to the myeloma cell line P3X63AG8.0653 using standard procedures. The resulting hybridomas are grown in bulk in HAT media for one week. Viable cells are recovered using lympholite M and cultured in HAT media plus cloning factor (Igen). After the hybridoma are grown for another week, a batch of the cells are cryopreserved in HAT media plus 10% DMSO. Another batch of

20        the cells is subdivided into individual clones by FACS sorting using the single cell deposit unit. The cells are sorted by forward and side scatter and for PI negative cells. The cells are grown up in HT media for two weeks. The supernatants are tested by Elisa and FACS (as described above) for monoclonal antibodies that can recognize native and recombinant MuSK-R protein. The antibodies are isotyped in an Elisa assay by using secondary antibodies that react with

25        IgG1, 2a, 2b, 3, IgM,  $\kappa$ , and  $\lambda$  (Caltag, CA).

30

Three monoclonal antibodies are identified H1, H2 and H4. All three can react with MuSK-R expressed on the cell line CEMSS-MuSK-R in a FACS assay. H1 is an IgG1,  $\kappa$ ; H2 is IgG1,  $\kappa$ ; H4 is IgM antibody.

cctaagatcc cgtccatcgc cactgggatg gtggggggccc tcctcttget gctgggtggtg 1980  
 gccctgggga tcggcctctt catgcgaagg cgccacatcg ttcggaagcg cacgctgcgg 2040  
 aggctgctgc aggagaggga gcttgtggag cctcttacac ccagtggaga agctcccaac 2100  
 caagctctct tgaggatctt gaaggaaact gaattcaaaa agatcaaagt gctgggctcc 2160  
 ggtgctgtcg gcacgggtgta taagggactc tggatcccag aaggtgagaa agttaaatt 2220  
 cccgtcgcta tcaaggaatt aagagaagca acatctccga aagccaaca ggaaatcctc 2280  
 gatgaagcct acgtgatggc cagcgtggac aacccccacg tgtgccgctt gctgggcatc 2340  
 tgcctcacct ccaccgtgca actcatcacg cagctcatgc ccttcggctg cctcctggac 2400  
 tatgtccggg aacacaaaaga caatattggc tcccagtacc tgcactaactg gtgtgtgcag 2460  
 atcgcaaagg gcatgaacta cttggaggac cgtcgcttgg tgcaccgcga cctggcagcc 2520  
 aggaacgtac tggtgaaaac accgcagcat gtcaagatca cagatttttg gctggccaaa 2580  
 ctgctgggtg cggaagagaa agaataccat gcagaaggag gcaaagtgcc tatcaagtgg 2640  
 atggcattgg aatcaatttt acacagaatc tataccacc agagtgatgt ctggagctac 2700  
 ggggtgaccg tttgggagtt gatgaccttt ggatccaagc catatgacgg aatccctgcc 2760  
 agcgagatct cctccatcct ggagaaagga gaacgcctcc ctcagccacc catatgtacc 2820  
 atcgatgtct acatgatcat ggtcaagtgc tggatgatag acgcagatag tcgccccaaag 2880  
 ttccgtgagt tgatcatcga attctccaaa atggcccag acccccagcg ctacctgtc 2940  
 attcaggggg atgaaagaat gcatttgcca agtcctacag actccaactt ctaccgtgcc 3000  
 ctgatggatg aagaagacat ggacgacgtg gtggatgccg acgagtacct catcccacag 3060  
 cagggtctct tcagcagccc ctccacgtca cggactcccc tcctgagctc tctgagtga 3120  
 accagcaaca attccaccgt ggcttgcatg gatagaaatg ggctgcaaag ctgtcccatc 3180  
 aaggaagaca gcttcttgca gcgatacagc tcagacccca caggcgctt gactgaggac 3240  
 agcatagacg acaccttctt cccagtgcct gaatacataa accagtccgt tcccaaaagg 3300  
 cccgtgggtc ctgtgcagaa tcctgtctat cacaatcagc ctctgaacct cgcgcccagc 3360  
 agagaccac actaccagga cccccacagc actgcagtgg gcaaccccg gatatctaac 3420  
 actgtccagc ccacctgtgt caacagcaca ttcgacagcc ctgcccactg ggcccagaaa 3480  
 ggcagccacc aaattagcct ggacaaccct gactaccagc aggacttctt tcccaaggaa 3540  
 gccaaagcaa atggcatctt taagggtccc acagctgaaa atgcagaata cctaagggtc 3600  
 gcgccacaaa gcagtgaatt tattggagca tga 3633

30

<210> 2  
 <211> 1210  
 <212> PRT  
 <213> EGFR  
 <400> 2

40

Met Arg Pro Ser Gly Thr Ala Gly Ala Ala Leu Leu Ala Leu Leu Ala  
 1 5 10 15  
 Ala Leu Cys Pro Ala Ser Arg Ala Leu Glu Glu Lys Lys Val Cys Gln  
 20 25 30  
 Gly Thr Ser Asn Lys Leu Thr Gln Leu Gly Thr Phe Glu Asp His Phe  
 35 40 45  
 Leu Ser Leu Gln Arg Met Phe Asn Asn Cys Glu Val Val Leu Gly Asn  
 50 55 60  
 Leu Glu Ile Thr Tyr Val Gln Arg Asn Tyr Asp Leu Ser Phe Leu Lys  
 65 70 75 80  
 Thr Ile Gln Glu Val Ala Gly Tyr Val Leu Ile Ala Leu Asn Thr Val  
 85 90 95

50

Glu Arg Ile Pro Leu Glu Asn Leu Gln Ile Ile Arg Gly Asn Met Tyr  
 100 105 110  
 Tyr Glu Asn Ser Tyr Ala Leu Ala Val Leu Ser Asn Tyr Asp Ala Asn  
 115 120 125  
 Lys Thr Gly Leu Lys Glu Leu Pro Met Arg Asn Leu Gln Glu Ile Leu  
 130 135 140

	His	Gly	Ala	Val	Arg	Phe	Ser	Asn	Asn	Pro	Ala	Leu	Cys	Asn	Val	Glu
	145					150					155					160
	Ser	Ile	Gln	Trp	Arg	Asp	Ile	Val	Ser	Ser	Asp	Phe	Leu	Ser	Asn	Met
					165					170					175	
	Ser	Met	Asp	Phe	Gln	Asn	His	Leu	Gly	Ser	Cys	Gln	Lys	Cys	Asp	Pro
			180						185					190		
	Ser	Cys	Pro	Asn	Gly	Ser	Cys	Trp	Gly	Ala	Gly	Glu	Glu	Asn	Cys	Gln
		195						200					205			
10	Lys	Leu	Thr	Lys	Ile	Ile	Cys	Ala	Gln	Gln	Cys	Ser	Gly	Arg	Cys	Arg
	210						215						220			
	Gly	Lys	Ser	Pro	Ser	Asp	Cys	Cys	His	Asn	Gln	Cys	Ala	Ala	Gly	Cys
	225					230					235					240
	Thr	Gly	Pro	Arg	Glu	Ser	Asp	Cys	Leu	Val	Cys	Arg	Lys	Phe	Arg	Asp
					245					250					255	
	Glu	Ala	Thr	Cys	Lys	Asp	Thr	Cys	Pro	Pro	Leu	Met	Leu	Tyr	Asn	Pro
			260					265						270		
	Thr	Thr	Tyr	Gln	Met	Asp	Val	Asn	Pro	Glu	Gly	Lys	Tyr	Ser	Phe	Gly
		275						280					285			
20	Ala	Thr	Cys	Val	Lys	Lys	Cys	Pro	Arg	Asn	Tyr	Val	Val	Thr	Asp	His
	290						295					300				
	Gly	Ser	Cys	Val	Arg	Ala	Cys	Gly	Ala	Asp	Ser	Tyr	Glu	Met	Glu	Glu
	305					310					315					320
	Asp	Gly	Val	Arg	Lys	Cys	Lys	Lys	Cys	Glu	Gly	Pro	Cys	Arg	Lys	Val
					325					330					335	
	Cys	Asn	Gly	Ile	Gly	Ile	Gly	Glu	Phe	Lys	Asp	Ser	Leu	Ser	Ile	Asn
			340						345					350		
	Ala	Thr	Asn	Ile	Lys	His	Phe	Lys	Asn	Cys	Thr	Ser	Ile	Ser	Gly	Asp
		355						360					365			
30	Leu	His	Ile	Leu	Pro	Val	Ala	Phe	Arg	Gly	Asp	Ser	Phe	Thr	His	Thr
	370						375					380				
	Pro	Pro	Leu	Asp	Pro	Gln	Glu	Leu	Asp	Ile	Leu	Lys	Thr	Val	Lys	Glu
	385					390					395					400
	Ile	Thr	Gly	Phe	Leu	Leu	Ile	Gln	Ala	Trp	Pro	Glu	Asn	Arg	Thr	Asp
					405				410						415	
	Leu	His	Ala	Phe	Glu	Asn	Leu	Glu	Ile	Ile	Arg	Gly	Arg	Thr	Lys	Gln
			420						425					430		
	His	Gly	Gln	Phe	Ser	Leu	Ala	Val	Val	Ser	Leu	Asn	Ile	Thr	Ser	Leu
		435						440					445			
40	Gly	Leu	Arg	Ser	Leu	Lys	Glu	Ile	Ser	Asp	Gly	Asp	Val	Ile	Ile	Ser
	450						455					460				
	Gly	Asn	Lys	Asn	Leu	Cys	Tyr	Ala	Asn	Thr	Ile	Asn	Trp	Lys	Lys	Leu
	465					470					475					480
	Phe	Gly	Thr	Ser	Gly	Gln	Lys	Thr	Lys	Ile	Ile	Ser	Asn	Arg	Gly	Glu
					485					490					495	
	Asn	Ser	Cys	Lys	Ala	Thr	Gly	Gln	Val	Cys	His	Ala	Leu	Cys	Ser	Pro
			500						505					510		
	Glu	Gly	Cys	Trp	Gly	Pro	Glu	Pro	Arg	Asp	Cys	Val	Ser	Cys	Arg	Asn
		515						520					525			
50	Val	Ser	Arg	Gly	Arg	Glu	Cys	Val	Asp	Lys	Cys	Asn	Leu	Leu	Glu	Gly
	530						535					540				
	Glu	Pro	Arg	Glu	Phe	Val	Glu	Asn	Ser	Glu	Cys	Ile	Gln	Cys	His	Pro
	545					550					555					560
	Glu	Cys	Leu	Pro	Gln	Ala	Met	Asn	Ile	Thr	Cys	Thr	Gly	Arg	Gly	Pro
					565					570					575	

	Asp	Asn	Cys	Ile	Gln	Cys	Ala	His	Tyr	Ile	Asp	Gly	Pro	His	Cys	Val
				580					585					590		
	Lys	Thr	Cys	Pro	Ala	Gly	Val	Met	Gly	Glu	Asn	Asn	Thr	Leu	Val	Trp
			595					600					605			
	Lys	Tyr	Ala	Asp	Ala	Gly	His	Val	Cys	His	Leu	Cys	His	Pro	Asn	Cys
		610					615					620				
	Thr	Tyr	Gly	Cys	Thr	Gly	Pro	Gly	Leu	Glu	Gly	Cys	Pro	Thr	Asn	Gly
	625					630					635					640
10	Pro	Lys	Ile	Pro	Ser	Ile	Ala	Thr	Gly	Met	Val	Gly	Ala	Leu	Leu	Leu
					645					650					655	
	Leu	Leu	Val	Val	Ala	Leu	Gly	Ile	Gly	Leu	Phe	Met	Arg	Arg	Arg	His
				660					665				670			
	Ile	Val	Arg	Lys	Arg	Thr	Leu	Arg	Arg	Leu	Leu	Gln	Glu	Arg	Glu	Leu
			675					680				685				
	Val	Glu	Pro	Leu	Thr	Pro	Ser	Gly	Glu	Ala	Pro	Asn	Gln	Ala	Leu	Leu
		690					695					700				
	Arg	Ile	Leu	Lys	Glu	Thr	Glu	Phe	Lys	Lys	Ile	Lys	Val	Leu	Gly	Ser
	705					710					715					720
20	Gly	Ala	Phe	Gly	Thr	Val	Tyr	Lys	Gly	Leu	Trp	Ile	Pro	Glu	Gly	Glu
					725					730					735	
	Lys	Val	Lys	Ile	Pro	Val	Ala	Ile	Lys	Glu	Leu	Arg	Glu	Ala	Thr	Ser
				740					745				750			
	Pro	Lys	Ala	Asn	Lys	Glu	Ile	Leu	Asp	Glu	Ala	Tyr	Val	Met	Ala	Ser
			755					760				765				
	Val	Asp	Asn	Pro	His	Val	Cys	Arg	Leu	Leu	Gly	Ile	Cys	Leu	Thr	Ser
		770					775					780				
	Thr	Val	Gln	Leu	Ile	Thr	Gln	Leu	Met	Pro	Phe	Gly	Cys	Leu	Leu	Asp
	785					790					795					800
30	Tyr	Val	Arg	Glu	His	Lys	Asp	Asn	Ile	Gly	Ser	Gln	Tyr	Leu	Leu	Asn
					805					810					815	
	Trp	Cys	Val	Gln	Ile	Ala	Lys	Gly	Met	Asn	Tyr	Leu	Glu	Asp	Arg	Arg
				820					825					830		
	Leu	Val	His	Arg	Asp	Leu	Ala	Ala	Arg	Asn	Val	Leu	Val	Lys	Thr	Pro
			835					840					845			
	Gln	His	Val	Lys	Ile	Thr	Asp	Phe	Gly	Leu	Ala	Lys	Leu	Leu	Gly	Ala
							855					860				
	Glu	Glu	Lys	Glu	Tyr	His	Ala	Glu	Gly	Gly	Lys	Val	Pro	Ile	Lys	Trp
	865					870					875					880
40	Met	Ala	Leu	Glu	Ser	Ile	Leu	His	Arg	Ile	Tyr	Thr	His	Gln	Ser	Asp
					885					890					895	
	Val	Trp	Ser	Tyr	Gly	Val	Thr	Val	Trp	Glu	Leu	Met	Thr	Phe	Gly	Ser
				900					905					910		
	Lys	Pro	Tyr	Asp	Gly	Ile	Pro	Ala	Ser	Glu	Ile	Ser	Ser	Ile	Leu	Glu
			915					920					925			
	Lys	Gly	Glu	Arg	Leu	Pro	Gln	Pro	Pro	Ile	Cys	Thr	Ile	Asp	Val	Tyr
		930					935					940				
	Met	Ile	Met	Val	Lys	Cys	Trp	Met	Ile	Asp	Ala	Asp	Ser	Arg	Pro	Lys
	945					950					955					960
50	Phe	Arg	Glu	Leu	Ile	Ile	Glu	Phe	Ser	Lys	Met	Ala	Arg	Asp	Pro	Gln
					965					970					975	
	Arg	Tyr	Leu	Val	Ile	Gln	Gly	Asp	Glu	Arg	Met	His	Leu	Pro	Ser	Pro
				980					985					990		
	Thr	Asp	Ser	Asn	Phe	Tyr	Arg	Ala	Leu	Met	Asp	Glu	Glu	Asp	Met	Asp
			995					1000						1005		

10	Asp Val Val Asp Ala Asp Glu Tyr Leu Ile Pro Gln Gln Gly Phe	
	1010 1015 1020	
	Phe Ser Ser Pro Ser Thr Ser Arg Thr Pro Leu Leu Ser Ser Leu	
	1025 1030 1035	
	Ser Ala Thr Ser Asn Asn Ser Thr Val Ala Cys Ile Asp Arg Asn	
	1040 1045 1050	
	Gly Leu Gln Ser Cys Pro Ile Lys Glu Asp Ser Phe Leu Gln Arg	
	1055 1060 1065	
	Tyr Ser Ser Asp Pro Thr Gly Ala Leu Thr Glu Asp Ser Ile Asp	
	1070 1075 1080	
20	Asp Thr Phe Leu Pro Val Pro Glu Tyr Ile Asn Gln Ser Val Pro	
	1085 1090 1095	
	Lys Arg Pro Ala Gly Ser Val Gln Asn Pro Val Tyr His Asn Gln	
	1100 1105 1110	
	Pro Leu Asn Pro Ala Pro Ser Arg Asp Pro His Tyr Gln Asp Pro	
	1115 1120 1125	
	His Ser Thr Ala Val Gly Asn Pro Glu Tyr Leu Asn Thr Val Gln	
	1130 1135 1140	
	Pro Thr Cys Val Asn Ser Thr Phe Asp Ser Pro Ala His Trp Ala	
	1145 1150 1155	
30	Gln Lys Gly Ser His Gln Ile Ser Leu Asp Asn Pro Asp Tyr Gln	
	1160 1165 1170	
	Gln Asp Phe Phe Pro Lys Glu Ala Lys Pro Asn Gly Ile Phe Lys	
	1175 1180 1185	
	Gly Ser Thr Ala Glu Asn Ala Glu Tyr Leu Arg Val Ala Pro Gln	
	1190 1195 1200	
	Ser Ser Glu Phe Ile Gly Ala	
	1205 1210	
	<210> 3	
	<211> 31	
	<212> DNA	
	<213> primer	
	<400> 3	
	ctaggctagc atgcgaccct ccgggacggc c	31
40	<210> 4	
	<211> 42	
	<212> DNA	
	<213> primer	
	<400> 4	
	ctctgcccgg cgactcgggc tgacagctat gagatggagg aa	42
50	<210> 5	
	<211> 39	
	<212> DNA	
	<213> primer	
	<400> 5	
	ttcctccatc tcatagctag cccgactcgc cgggcagag	39

<210> 6  
 <211> 31  
 <212> DNA  
 <213> primer  
 <400> 6  
 ggatataccta cgtgcgcttc cgaacgatgt g

31

10

<210> 7  
 <211> 2607  
 <212> DNA  
 <213> MuSK  
 <400> 7

20

30

40

50

atgagagagc	tcgtcaacat	tccactggta	catattctta	ctctgggttg	cttcagcgga	60
actgagaaac	ttccaaaagc	tcctgtcatc	accactcctc	ttgaaacagt	ggatgcctta	120
ggtgaagaag	tggctacttt	catgtgtgca	gtggaatcct	acccccagcc	tgagatttcc	180
tggactagaa	ataaaaattct	cattaaactc	tttgacaccc	ggtacagcat	ccgggagaat	240
gggcagctcc	tcaccatcct	gagtgtggaa	gacagtgatg	atggcattta	ctgctgcacg	300
gccaacaatg	gtgtggggagg	agctgtggag	agttgtggag	ccctgcaagt	gaagatgaaa	360
cctaaaataa	ctcgtcctcc	cataaatgtg	aaaataatag	agggattaaa	agcagtccta	420
ccatgtacta	caatgggtaa	tcccaaacca	tcagtgtctt	ggataaaggg	agacagccct	480
ctcagggaaa	attcccgaat	tgcagttctt	gaatctggga	gcttgaggat	tcataacgta	540
caaaaggaag	atgcaggaca	gtatcgatgt	gtggcaaaaa	acagcctcgg	gacagcatat	600
tccaaagtgg	tgaagctgga	at ttgaggtt	tttgccagga	tcctgcgggc	tcctgaatcc	660
cacaatgtca	cctttggctc	ctttgtgacc	ctgcactgta	cagcaacagg	cattcctgtc	720
cccaccatca	cctggattga	aaacggaaat	gctgtttctt	ctgggtccat	tcaagagagt	780
gtgaaagacc	gagtgattga	ctcaagactg	cagctgttta	tcaccaagcc	aggactctac	840
acatgcatag	ctaccaataa	gcatggggag	aagttcagta	ctgccaaagg	tgagccacc	900
atcagcatag	cagaatggag	taaaccacag	aaagataaca	aaggctactg	cgcccagtag	960
agagggggagg	tgtgtaatgc	agtcctggca	aaagatgctc	ttgtttttct	caacacctcc	1020
tatgcggaac	ctgaggaggc	ccaagagcta	ctgggtccaca	cggcctggaa	tgaactgaaa	1080
gtagtgaacc	cagctgccc	gccagctgct	gaggctttgt	tgtgtaacca	catcttccag	1140
gagtgcagtc	ctggagtagt	gcctactcct	attcccattt	gcagagagta	ctgcttggca	1200
gtaaaggagc	tcttctgcgc	aaaagaatgg	ctggtaatgg	aagagaagac	ccacagagga	1260
ctctacagat	ccgagatgca	tttgctgtcc	gtgccaaaat	gcagcaagct	tcccagcatg	1320
cattgggacc	ccacggcctg	tgccagactg	ccacatctag	attataacaa	agaaaacctt	1380
aaaacattcc	caccaatgac	gtcctcaaag	ccaagtgtgg	acattccaaa	tctgccttcc	1440
tcctcctctt	cttccttctc	tgtctcacct	acatactcca	tgactgtaat	aatctccatc	1500
atgtccagct	ttgcaatatt	tgtgcttctt	accataacta	ctctctattg	ctgccgaaga	1560
agaaaacaat	ggaaaaataa	gaaaagagaa	tcagcagcag	taaccctcac	cacactgcct	1620
tctgagctct	tactagatag	acttcatccc	aaccccatgt	accagaggat	gccgctcctt	1680
ctgaacccca	aattgctcag	cctggagtat	ccaaggaata	acattgaata	tgtgagagac	1740
atcggagagg	gagcgttttg	aagggtgttt	caagcaaggg	caccaggctt	acttccctat	1800
gaacctttca	ctatgggtggc	agtaaagatg	ctcaaagaag	aagcctcggc	agatatgcaa	1860
gcggactttc	agagggaggc	agccctcatg	gcagaatttg	acaaccctaa	cattgtgaag	1920
ctattaggag	tgtgtgctgt	cggaagcca	atgtgcctgc	tctttgaata	catggcctat	1980
ggtgacctca	atgagttcct	ccgcagcatg	tcccctcaca	ccgtgtgcag	cctcagtcac	2040
agtgacttgt	ctatgagggc	tcaggctctc	agccctgggc	ccccacccct	ctcctgtgct	2100
gagcagcttt	gcattgccag	gcagggtggca	gctggcatgg	cttacctctc	agaacgtaag	2160
tttgttcacc	gagatttagc	caccaggaac	tgcttgggtg	gcgagaacat	ggtggtgaaa	2220
attgccgact	ttggcctctc	caggaacatc	tactcagcag	actactacaa	agctaataaa	2280
aacgacgcta	tccctatccg	ttggatgcca	ccagagtcca	ttttttataa	ccgctacact	2340
acagagtctg	atgtgtgggc	ctatggcggtg	gtcctctggg	agatcttctc	ctatggcctg	2400
cagccctact	atgggatggc	ccatgaggag	gtcatttact	acgtgcgaga	tggcaacatc	2460

ctctcctgcc ctgagaactg ccccgtaggag ctgtacaatc tcatgcgtct atgttggagc 2520  
aagctgcctg cagacagacc cagtttcacc agtattcacc gaattctgga acgcatgtgt 2580  
gagagggcag agggaactgt gagtgtc 2607

<210> 8

<211> 869

<212> PRT

<213> MuSK

<400> 8

10

Met Arg Glu Leu Val Asn Ile Pro Leu Val His Ile Leu Thr Leu Val  
1 5 10 15

Ala Phe Ser Gly Thr Glu Lys Leu Pro Lys Ala Pro Val Ile Thr Thr  
20 25 30

Pro Leu Glu Thr Val Asp Ala Leu Val Glu Glu Val Ala Thr Phe Met  
35 40 45

Cys Ala Val Glu Ser Tyr Pro Gln Pro Glu Ile Ser Trp Thr Arg Asn  
50 55 60

20

Lys Ile Leu Ile Lys Leu Phe Asp Thr Arg Tyr Ser Ile Arg Glu Asn  
65 70 75 80

Gly Gln Leu Leu Thr Ile Leu Ser Val Glu Asp Ser Asp Asp Gly Ile  
85 90 95

Tyr Cys Cys Thr Ala Asn Asn Gly Val Gly Gly Ala Val Glu Ser Cys  
100 105 110

Gly Ala Leu Gln Val Lys Met Lys Pro Lys Ile Thr Arg Pro Pro Ile  
115 120 125

Asn Val Lys Ile Ile Glu Gly Leu Lys Ala Val Leu Pro Cys Thr Thr  
130 135 140

30

Met Gly Asn Pro Lys Pro Ser Val Ser Trp Ile Lys Gly Asp Ser Pro  
145 150 155 160

Leu Arg Glu Asn Ser Arg Ile Ala Val Leu Glu Ser Gly Ser Leu Arg  
165 170 175

Ile His Asn Val Gln Lys Glu Asp Ala Gly Gln Tyr Arg Cys Val Ala  
180 185 190

Lys Asn Ser Leu Gly Thr Ala Tyr Ser Lys Val Val Lys Leu Glu Phe  
195 200 205

Glu Val Phe Ala Arg Ile Leu Arg Ala Pro Glu Ser His Asn Val Thr  
210 215 220

40

Phe Gly Ser Phe Val Thr Leu His Cys Thr Ala Thr Gly Ile Pro Val  
225 230 235 240

Pro Thr Ile Thr Trp Ile Glu Asn Gly Asn Ala Val Ser Ser Gly Ser  
245 250 255

Ile Gln Glu Ser Val Lys Asp Arg Val Ile Asp Ser Arg Leu Gln Leu  
260 265 270

Phe Ile Thr Lys Pro Gly Leu Tyr Thr Cys Ile Ala Thr Asn Lys His  
275 280 285

Gly Glu Lys Phe Ser Thr Ala Lys Ala Ala Ala Thr Ile Ser Ile Ala  
290 295 300

50

Glu Trp Ser Lys Pro Gln Lys Asp Asn Lys Gly Tyr Cys Ala Gln Tyr  
305 310 315 320

Arg Gly Glu Val Cys Asn Ala Val Leu Ala Lys Asp Ala Leu Val Phe  
325 330 335

Leu Asn Thr Ser Tyr Ala Asp Pro Glu Glu Ala Gln Glu Leu Leu Val  
340 345 350



	His	Thr	Ala	Trp	Asn	Glu	Leu	Lys	Val	Val	Ser	Pro	Val	Cys	Arg	Pro	
			355					360					365				
	Ala	Ala	Glu	Ala	Leu	Leu	Cys	Asn	His	Ile	Phe	Gln	Glu	Cys	Ser	Pro	
			370				375					380					
	Gly	Val	Val	Pro	Thr	Pro	Ile	Pro	Ile	Cys	Arg	Glu	Tyr	Cys	Leu	Ala	
						390					395					400	
	Val	Lys	Glu	Leu	Phe	Cys	Ala	Lys	Glu	Trp	Leu	Val	Met	Glu	Glu	Lys	
					405					410					415		
10	Thr	His	Arg	Gly	Leu	Tyr	Arg	Ser	Glu	Met	His	Leu	Leu	Ser	Val	Pro	
				420					425					430			
	Lys	Cys	Ser	Lys	Leu	Pro	Ser	Met	His	Trp	Asp	Pro	Thr	Ala	Cys	Ala	
			435					440					445				
	Arg	Leu	Pro	His	Leu	Asp	Tyr	Asn	Lys	Glu	Asn	Leu	Lys	Thr	Phe	Pro	
			450				455					460					
	Pro	Met	Thr	Ser	Ser	Lys	Pro	Ser	Val	Asp	Ile	Pro	Asn	Leu	Pro	Ser	
						470				475					480		
	Ser	Ser	Ser	Ser	Ser	Phe	Ser	Val	Ser	Pro	Thr	Tyr	Ser	Met	Thr	Val	
					485					490					495		
20	Ile	Ile	Ser	Ile	Met	Ser	Ser	Phe	Ala	Ile	Phe	Val	Leu	Leu	Thr	Ile	
				500					505						510		
	Thr	Thr	Leu	Tyr	Cys	Cys	Arg	Arg	Arg	Lys	Gln	Trp	Lys	Asn	Lys	Lys	
			515					520					525				
	Arg	Glu	Ser	Ala	Ala	Val	Thr	Leu	Thr	Thr	Leu	Pro	Ser	Glu	Leu	Leu	
			530				535					540					
	Leu	Asp	Arg	Leu	His	Pro	Asn	Pro	Met	Tyr	Gln	Arg	Met	Pro	Leu	Leu	
						550				555					560		
	Leu	Asn	Pro	Lys	Leu	Leu	Ser	Leu	Glu	Tyr	Pro	Arg	Asn	Asn	Ile	Glu	
				565						570					575		
30	Tyr	Val	Arg	Asp	Ile	Gly	Glu	Gly	Ala	Phe	Gly	Arg	Val	Phe	Gln	Ala	
				580				585						590			
	Arg	Ala	Pro	Gly	Leu	Leu	Pro	Tyr	Glu	Pro	Phe	Thr	Met	Val	Ala	Val	
			595					600					605				
	Lys	Met	Leu	Lys	Glu	Glu	Ala	Ser	Ala	Asp	Met	Gln	Ala	Asp	Phe	Gln	
			610				615					620					
	Arg	Glu	Ala	Ala	Leu	Met	Ala	Glu	Phe	Asp	Asn	Pro	Asn	Ile	Val	Lys	
						630				635					640		
	Leu	Leu	Gly	Val	Cys	Ala	Val	Gly	Lys	Pro	Met	Cys	Leu	Leu	Phe	Glu	
					645					650					655		
40	Tyr	Met	Ala	Tyr	Gly	Asp	Leu	Asn	Glu	Phe	Leu	Arg	Ser	Met	Ser	Pro	
				660				665						670			
	His	Thr	Val	Cys	Ser	Leu	Ser	His	Ser	Asp	Leu	Ser	Met	Arg	Ala	Gln	
			675					680					685				
	Val	Ser	Ser	Pro	Gly	Pro	Pro	Pro	Leu	Ser	Cys	Ala	Glu	Gln	Leu	Cys	
			690				695					700					
	Ile	Ala	Arg	Gln	Val	Ala	Ala	Gly	Met	Ala	Tyr	Leu	Ser	Glu	Arg	Lys	
						710				715					720		
	Phe	Val	His	Arg	Asp	Leu	Ala	Thr	Arg	Asn	Cys	Leu	Val	Gly	Glu	Asn	
					725					730					735		
50	Met	Val	Val	Lys	Ile	Ala	Asp	Phe	Gly	Leu	Ser	Arg	Asn	Ile	Tyr	Ser	
				740				745						750			
	Ala	Asp	Tyr	Tyr	Lys	Ala	Asn	Glu	Asn	Asp	Ala	Ile	Pro	Ile	Arg	Trp	
			755					760					765				
	Met	Pro	Pro	Glu	Ser	Ile	Phe	Tyr	Asn	Arg	Tyr	Thr	Thr	Glu	Ser	Asp	
				770				775					780				

	Val Trp Ala Tyr Gly Val Val Leu Trp Glu Ile Phe Ser Tyr Gly Leu	
	785 790 795 800	
	Gln Pro Tyr Tyr Gly Met Ala His Glu Glu Val Ile Tyr Tyr Val Arg	
	805 810 815	
	Asp Gly Asn Ile Leu Ser Cys Pro Glu Asn Cys Pro Val Glu Leu Tyr	
	820 825 830	
	Asn Leu Met Arg Leu Cys Trp Ser Lys Leu Pro Ala Asp Arg Pro Ser	
	835 840 845	
10	Phe Thr Ser Ile His Arg Ile Leu Glu Arg Met Cys Glu Arg Ala Glu	
	850 855 860	
	Gly Thr Val Ser Val	
	865	
	<210> 9	
	<211> 25	
	<212> DNA	
	<213> primer	
20	<400> 9	
	cgtcctgcgt gagcctggat taatc	25
	<210> 10	
	<211> 25	
	<212> DNA	
	<213> primer	
	<400> 10	
	gcctggatta atcatgagag agctc	25
30	<210> 11	
	<211> 42	
	<212> DNA	
	<213> primer	
	<400> 11	
	cgaggcctgt cttcaacctt agacactcac agttccctct gc	42
40	<210> 12	
	<211> 27	
	<212> DNA	
	<213> primer	
	<400> 12	
	cggcctgtgc cagactgccca catctag	27
50	<210> 13	
	<211> 32	
	<212> DNA	
	<213> primer	
	<400> 13	
	cgtctaggtg agggttactg ctgctgattc tc	32

<210> 14  
 <211> 38  
 <212> DNA  
 <213> primer  
 <400> 14  
 ggттаaccct attcaatggt attccttgga tactccag 38

10

<210> 15  
 <211> 24  
 <212> DNA  
 <213> primer  
 <400> 15  
 cttccaaaag ctctgtcat cacc 24

20

<210> 16  
 <211> 26  
 <212> DNA  
 <213> primer  
 <400> 16  
 ccagtcattgg agtatgtagg tgagac 26

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100

## IT IS CLAIMED:

1. A method of identifying genetically modified mammalian cells comprising the steps of:
  - a) introducing a nucleic acid sequence encoding a mutated protein-tyrosine kinase receptor (PTKR) operatively linked to an expression control sequence into a cell to form a genetically modified cell, wherein the mutated PTKR either comprises a modification to the intracellular and the extracellular domains, or excludes any nerve growth factor receptor(s) (NGFR) and comprises a modification to the intracellular domain;
  - b) allowing expression of the mutated PTKR in the genetically modified cell; and
  - 10 c) identifying said genetically modified cell expressing the mutated PTKR.
2. The method according to claim 1, wherein the mutated PTKR is selected from a mutated epidermal growth factor receptor (EGFR) family member and muscle specific tyrosine kinase receptor (MuSK-R) family member.
3. The method according to claim 2, wherein the mutated EGFR family member is a mutated EGFR1, optionally selected from the sequence designated EGFR1-I and EGFR1-II.
4. The method according to claim 2, wherein the mutated MuSK-R is selected from the
  - 20 sequence designated mMuSK-RI and mMuSK-RII.
5. The method according to any preceding claims, wherein the mutated PTKR is truncated from the intracellular domain, and optionally also from the extracellular domain.
6. The method according to claim 2, wherein the introducing step is accomplished by incorporating the nucleic acid sequence encoding the mutated EGFR or MuSK-R into a vector and introducing said vector into said cell.
7. The method according to claim 6, wherein the vector is a retroviral vector.

8. The method according to claim 1, wherein said identifying step is accomplished by contacting the genetically modified cells with an antibody that recognizes and binds to the mutated PTKR.

9. The method according to claim 1, wherein the identifying step separates the genetically modified cells from the non-genetically modified cells.

10. The method according to claim 1, further comprising the step of separating the identified cells expressing the mutated PTKR.

10

11. The method according to claim 1, wherein the cells are human cells.

12. The method according to claim 11, wherein the cells are selected from the group consisting of hematopoietic cells, liver cells, endothelial cells and smooth muscle cells.

13. The method according to claim 11, wherein the cells are hematopoietic cells.

14. The method according to claim 13, wherein the cells are hematopoietic cells are stem cells or T-cells.

20

15. The method according to claim 6, wherein a heterologous gene is also incorporated into said vector.

16. A method of identifying genetically modified mammalian cells comprising the steps of:

- a) incorporating into a vector a nucleic acid sequence encoding a mutated protein-tyrosine kinase receptor (PTKR) family member, wherein said PTKR either comprises a modification to the intracellular and the extracellular domains, or excludes any nerve growth factor receptor(s) (NGFR) and comprises a modification to the intracellular domain; introducing the vector into a mammalian cell to form a genetically modified cell;

30

- b) allowing expression of the mutated PTKR in the genetically modified cell; and

c) identifying said genetically modified cell expressing the mutated PTKR.

17. The method according to claim 16, wherein the mutated PTKR is selected from EGFR, preferably EGFR1-I or EGFR1-II, and MuSK-R, preferably mMuSK-RI or mMuSK-RII.

18. The method according to claim 16, wherein the vector is a retroviral vector.

19. The method according to claim 16, wherein a heterologous gene is also incorporated into said vector.

10

20. A method for the immunoselection of transduced mammalian cells comprising,

- a) retrovirally transducing mammalian cells with a nucleic acid sequence encoding a protein-tyrosine kinase receptor (PTKR) family member operatively linked to an expression control sequence; wherein said PTKR either comprises a modification to the intracellular and the extracellular domains, or excludes any nerve growth factor receptor(s) (NGFR) and comprises a modification to the intracellular domain;
- b) incubating the transduced cells with a marked antibody which recognizes and binds specifically to the mutated PTKR; and
- c) identifying the marked transduced cells.

20

21. The method according to claim 20, wherein the cells are human cells or hematopoietic cells.

22. The method according to claim 20, wherein the mutated PTKR is selected from EGFR, preferably EGFR1-I or EGFR1-II, and MuSK-R, preferably mMuSK-RI or mMuSK-RII.

23. The method according to claim 20, wherein the cells are transduced by a retroviral vector derived from the group consisting of moloney murine leukemia virus (MoMLV), myeloproliferative sarcoma virus (MPSV), murine embryonic stem cell virus (MESV), murine stem cell virus (MSCV) and spleen focus forming virus (SFFV).

30

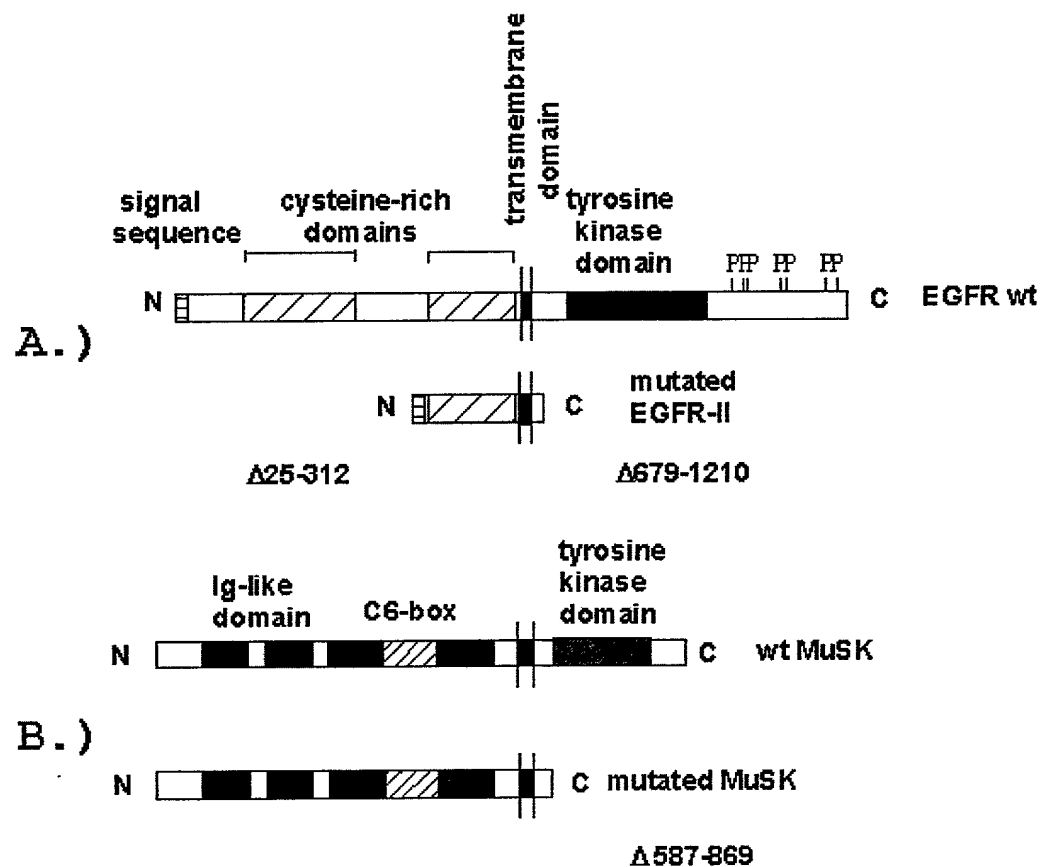
24. The method according to claim 20, wherein the cells are transduced by a vector derived from a lentivirus.
25. The method according to claim 20, further comprising the step of separating the identified marked transduced cells from non-marked cells.
26. The method according to claim 20, further comprising the step of expanding the marked transduced cells.
- 10 27. A method of identifying mammalian cells expressing a protein of interest, comprising the steps of,
  - a) introducing into a mammalian cell a nucleic acid encoding a nucleic acid comprising a DNA sequence encoding a protein of interest and encoding a protein-tyrosine kinase receptor (PTKR) family member operatively linked to an expression control sequence, wherein said PTKR either comprises a modification to the intracellular and the extracellular domains, or excludes any nerve growth factor receptor(s) (NGFR) and comprises a modification to the intracellular domain;
  - b) culturing the resulting mammalian cells; and
  - c) identifying cells which express the mutated PTKR thereby obtaining cells which
- 20 express the protein of interest.
28. The method according to claim 27, wherein the mutated PTKR is selected from EGFR, preferably EGFR1-I or EGFR1-II, and MuSK-R, preferably mMuSK-RI or mMuSK-RII.
29. The method according to claim 27, wherein the nucleic acid encoding the mutated PTKR and the nucleic acid encoding the protein of interest in step (a) are introduced on a retroviral vector.
30. A mutated muscle specific tyrosine kinase receptor (MuSK-R) family member which is truncated within the cytoplasmic domain, preferably mMuSK-RI or mMuSK-RII.
- 30

**ABSTRACT**

The present invention concerns a method of identifying genetically modified mammalian cells using a mutated protein-tyrosine kinase receptor (PTKR) as a selectable marker in mammalian cells. Particularly preferred mutated PTKR selective markers are mutated epidermal growth factor receptor (EGFR) family members, and muscle specific tyrosine kinase receptor (MuSK-R) family members. Further a method for the immunoselection of transduced mammalian cells is disclosed comprising retrovirally transducing mammalian cells with a nucleic acid sequence encoding a mutated EGFR, incubating the transduced cells with a marked antibody which  
 10 recognizes and binds specifically to the mutated PTKR, and identifying the marked transduced cells.

11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000  
1001  
1002  
1003  
1004  
1005  
1006  
1007  
1008  
1009  
1010  
1011  
1012  
1013  
1014  
1015  
1016  
1017  
1018  
1019  
1020  
1021  
1022  
1023  
1024  
1025  
1026  
1027  
1028  
1029  
1030  
1031  
1032  
1033  
1034  
1035  
1036  
1037  
1038  
1039  
1040  
1041  
1042  
1043  
1044  
1045  
1046  
1047  
1048  
1049  
1050  
1051  
1052  
1053  
1054  
1055  
1056  
1057  
1058  
1059  
1060  
1061  
1062  
1063  
1064  
1065  
1066  
1067  
1068  
1069  
1070  
1071  
1072  
1073  
1074  
1075  
1076  
1077  
1078  
1079  
1080  
1081  
1082  
1083  
1084  
1085  
1086  
1087  
1088  
1089  
1090  
1091  
1092  
1093  
1094  
1095  
1096  
1097  
1098  
1099  
1100  
1101  
1102  
1103  
1104  
1105  
1106  
1107  
1108  
1109  
1110  
1111  
1112  
1113  
1114  
1115  
1116  
1117  
1118  
1119  
1120  
1121  
1122  
1123  
1124  
1125  
1126  
1127  
1128  
1129  
1130  
1131  
1132  
1133  
1134  
1135  
1136  
1137  
1138  
1139  
1140  
1141  
1142  
1143  
1144  
1145  
1146  
1147  
1148  
1149  
1150  
1151  
1152  
1153  
1154  
1155  
1156  
1157  
1158  
1159  
1160  
1161  
1162  
1163  
1164  
1165  
1166  
1167  
1168  
1169  
1170  
1171  
1172  
1173  
1174  
1175  
1176  
1177  
1178  
1179  
1180  
1181  
1182  
1183  
1184  
1185  
1186  
1187  
1188  
1189  
1190  
1191  
1192  
1193  
1194  
1195  
1196  
1197  
1198  
1199  
1200  
1201  
1202  
1203  
1204  
1205  
1206  
1207  
1208  
1209  
1210  
1211  
1212  
1213  
1214  
1215  
1216  
1217  
1218  
1219  
1220  
1221  
1222  
1223  
1224  
1225  
1226  
1227  
1228  
1229  
1230  
1231  
1232  
1233  
1234  
1235  
1236  
1237  
1238  
1239  
1240  
1241  
1242  
1243  
1244  
1245  
1246  
1247  
1248  
1249  
1250  
1251  
1252  
1253  
1254  
1255  
1256  
1257  
1258  
1259  
1260  
1261  
1262  
1263  
1264  
1265  
1266  
1267  
1268  
1269  
1270  
1271  
1272  
1273  
1274  
1275  
1276  
1277  
1278  
1279  
1280  
1281  
1282  
1283  
1284  
1285  
1286  
1287  
1288  
1289  
1290  
1291  
1292  
1293  
1294  
1295  
1296  
1297  
1298  
1299  
1300  
1301  
1302  
1303  
1304  
1305  
1306  
1307  
1308  
1309  
1310  
1311  
1312  
1313  
1314  
1315  
1316  
1317  
1318  
1319  
1320  
1321  
1322  
1323  
1324  
1325  
1326  
1327  
1328  
1329  
1330  
1331  
1332  
1333  
1334  
1335  
1336  
1337  
1338  
1339  
1340  
1341  
1342  
1343  
1344  
1345  
1346  
1347  
1348  
1349  
1350  
1351  
1352  
1353  
1354  
1355  
1356  
1357  
1358  
1359  
1360  
1361  
1362  
1363  
1364  
1365  
1366  
1367  
1368  
1369  
1370  
1371  
1372  
1373  
1374  
1375  
1376  
1377  
1378  
1379  
1380  
1381  
1382  
1383  
1384  
1385  
1386  
1387  
1388  
1389  
1390  
1391  
1392  
1393  
1394  
1395  
1396  
1397  
1398  
1399  
1400  
1401  
1402  
1403  
1404  
1405  
1406  
1407  
1408  
1409  
1410  
1411  
1412  
1413  
1414  
1415  
1416  
1417  
1418  
1419  
1420  
1421  
1422  
1423  
1424  
1425  
1426  
1427  
1428  
1429  
1430  
1431  
1432  
1433  
1434  
1435  
1436  
1437  
1438  
1439  
1440  
1441  
1442  
1443  
1444  
1445  
1446  
1447  
1448  
1449  
1450  
1451  
1452  
1453  
1454  
1455  
1456  
1457  
1458  
1459  
1460  
1461  
1462  
1463  
1464  
1465  
1466  
1467  
1468  
1469  
1470  
1471  
1472  
1473  
1474  
1475  
1476  
1477  
1478  
1479  
1480  
1481  
1482  
1483  
1484  
1485  
1486  
1487  
1488  
1489  
1490  
1491  
1492  
1493  
1494  
1495  
1496  
1497  
1498  
1499  
1500  
1501  
1502  
1503  
1504  
1505  
1506  
1507  
1508  
1509  
1510  
1511  
1512  
1513  
1514  
1515  
1516  
1517  
1518  
1519  
1520  
1521  
1522  
1523  
1524  
1525  
1526  
1527  
1528  
1529  
1530  
1531  
1532  
1533  
1534  
1535  
1536  
1537  
1538  
1539  
1540  
1541  
1542  
1543  
1544  
1545  
1546  
1547  
1548  
1549  
1550  
1551  
1552  
1553  
1554  
1555  
1556  
1557  
1558  
1559  
1560  
1561  
1562  
1563  
1564  
1565  
1566  
1567  
1568  
1569  
1570  
1571  
1572  
1573  
1574  
1575  
1576  
1577  
1578  
1579  
1580  
1581  
1582  
1583  
1584  
1585  
1586  
1587  
1588  
1589  
1590  
1591  
1592  
1593  
1594  
1595  
1596  
1597  
1598  
1599  
1600  
1601  
1602  
1603  
1604  
1605  
1606  
1607  
1608  
1609  
1610  
1611  
1612  
1613  
1614  
1615  
1616  
1617  
1618  
1619  
1620  
1621  
1622  
1623  
1624  
1625  
1626  
1627  
1628  
1629  
1630  
1631  
1632  
1633  
1634  
1635  
1636  
1637  
1638  
1639  
1640  
1641  
1642  
1643  
1644  
1645  
1646  
1647  
1648  
1649  
1650  
1651  
1652  
1653  
1654  
1655  
1656  
1657  
1658  
1659  
1660  
1661  
1662  
1663  
1664  
1665  
1666  
1667  
1668  
1669  
1670  
1671  
1672  
1673  
1674  
1675  
1676  
1677  
1678  
1679  
1680  
1681  
1682  
1683  
1684  
1685  
1686  
1687  
1688  
1689  
1690  
1691  
1692  
1693  
1694  
1695  
1696  
1697  
1698  
1699  
1700  
1701  
1702  
1703  
1704  
1705  
1706  
1707  
1708  
1709  
1710  
1711  
1712  
1713  
1714  
1715  
1716  
1717  
1718  
1719  
1720  
1721  
1722  
1723  
1724  
1725  
1726  
1727  
1728  
1729  
1730  
1731  
1732  
1733  
1734  
1735  
1736  
1737  
1738  
1739  
1740  
1741  
1742  
1743  
1744  
1745  
1746  
1747  
1748  
1749  
1750  
1751  
1752  
1753  
1754  
1755  
1756  
1757  
1758  
1759  
1760  
1761  
1762  
1763  
1764  
1765  
1766  
1767  
1768  
1769  
1770  
1771  
1772  
1773  
1774  
1775  
1776  
1777  
1778  
1779  
1780  
1781  
1782  
1783  
1784  
1785  
1786  
1787  
1788  
1789  
1790  
1791  
1792  
1793  
1794  
1795  
1796  
1797  
1798  
1799  
1800  
1801  
1802  
1803  
1804  
1805  
1806  
1807  
1808  
1809  
1810  
1811  
1812  
1813  
1814  
1815  
1816  
1817  
1818  
1819  
1820  
1821  
1822  
1823  
1824  
1825  
1826  
1827  
1828  
1829  
1830  
1831  
1832  
1833  
1834  
1835  
1836  
1837  
1838  
1839  
1840  
1841  
1842  
1843  
1844  
1845  
1846  
1847  
1848  
1849  
1850  
1851  
1852  
1853  
1854  
1855  
1856  
1857  
1858  
1859  
1860  
1861  
1862  
1863  
1864  
1865  
1866  
1867  
1868  
1869  
1870  
1871  
1872  
1873  
1874  
1875  
1876  
1877  
1878  
1879  
1880  
1881  
1882  
1883  
1884  
1885  
1886  
1887  
1888  
1889  
1890  
1891  
1892  
1893  
1894  
1895  
1896  
1897  
1898  
1899  
1900  
1901  
1902  
1903  
1904  
1905  
1906  
1907  
1908  
1909  
1910  
1911  
1912  
1913  
1914  
1915  
1916  
1917  
1918  
1919  
1920  
1921  
1922  
1923  
1924  
1925  
1926  
1927  
1928  
1929  
1930  
1931  
1932  
1933  
1934  
1935  
1936  
1937  
1938  
1939  
1940  
1941  
1942  
1943  
1944  
1945  
1946  
1947  
1948  
1949  
1950  
1951  
1952  
1953  
1954  
1955  
1956  
1957  
1958  
1959  
1960  
1961  
1962  
1963  
1964  
1965  
1966  
1967  
1968  
1969  
1970  
1971  
1972  
1973  
1974  
1975  
1976  
1977  
1978  
1979  
1980  
1981  
1982  
1983  
1984  
1985  
1986  
1987  
1988  
1989  
1990  
1991  
1992  
1993  
1994  
1995  
1996  
1997  
1998  
1999  
2000  
2001  
2002  
2003  
2004  
2005  
2006  
2007  
2008  
2009  
2010  
2011  
2012  
2013  
2014  
2015  
2016  
2017  
2018  
2019  
2020  
2021  
2022  
2023  
2024  
2025  
2026  
2027  
2028  
2029  
2030  
2031  
2032  
2033  
2034  
2035  
2036  
2037  
2038  
2039  
2040  
2041  
2042  
2043  
2044  
2045  
2046  
2047  
2048  
2049  
2050  
2051  
2052  
2053  
2054  
2055  
2056  
2057  
2058  
2059  
2060  
2061  
2062  
2063  
2064  
2065  
2066  
2067  
2068  
2069  
2070  
2071  
2072  
2073  
2074  
2075  
2076  
2077  
2078  
2079  
2080  
2081  
2082  
2083  
2084  
2085  
2086  
2087  
2088  
2089  
2090  
2091  
2092  
2093  
2094  
2095  
2096  
2097  
2098  
2099  
2100  
2101  
2102  
2103  
2104  
2105  
2106  
2107  
2108  
2109  
2110  
2111  
2112  
2113  
2114  
2115  
2116  
2117  
2118  
2119  
2120  
2121  
2122  
2123  
2124  
2125  
2126  
2127  
2128  
2129  
2130  
2131  
2132  
2133  
2134  
2135  
2136  
2137  
2138  
2139  
2140  
2141  
2142  
2143  
2144  
2145  
2146  
2147  
2148  
2149  
2150  
2151  
2152  
2153  
2154  
2155  
2156  
2157  
2158  
2159  
2160  
2161  
2162  
2163  
2164  
2165  
2166  
2167  
2168  
2169  
2170  
2171  
2172  
2173  
2174  
2175  
2176  
2177  
2178  
2179  
2180  
2181  
2182  
2183  
2184  
2185  
2186  
2187  
2188  
2189  
2190  
2191  
2192  
2193  
2194  
2195  
2196  
2197  
2198  
2199  
2200  
2201  
2202  
2203  
2204  
2205  
2206  
2





In A.) the the wild-type and the mutated EGFR-II are shown, in B.) the wild type and the mutated MuSK is shown. Both EGFR-II and the mutated MuSK can be used as selectable cell surface markers.

FIGURE 1

	MRPSGTAGAA	LLALLAALCP	ASRALEEKKV	CQGTSNKLTQ	LGTFEDHFLS	50
	LQRMFNNEV	VLGNLEITYV	QRNYDLSFLK	TIQEVAGYVL	IALNTVERIP	100
	LENLQIIRGN	MYYENSYALA	VLSNYDANKT	GLKELPMRNL	QEILHGAVRF	150
	SNNPALCNVE	SIQWRDIVSS	DFLSNMSMDF	QNHLGSCQKC	DPSCPNGSCW	200
5	GAGEENCQKL	TKIICAQQCS	GRCRGKSPSD	CCHNQCAAGC	TGPRESDECLV	250
	CRKFRDEATC	KDTCPPMLY	NPTTYQMDVN	PEGKYSFGAT	CVKKCPRNYV	300
	VTDHGSCVRA	CGADSYEMEE	DGVRKCKKCE	GPCRKVCNGI	GIGEFKDSLS	350
	INATNIKHFK	NCTSIGDLH	ILPVAFRGDS	FTHTPPLDPQ	ELDILKTVKE	400
	ITGFLLIQAW	PENRTDLHAF	ENLEIIRGRT	KQHGQFSLAV	VSLNITSLGL	450
10	RSLKEISDGD	VIISGNKNLC	YANTINWKKL	FGTSGQKTKI	ISNRGENSCK	500
	ATGQVCHALC	SPEGCWGPEP	RDCVSCRNVS	RGRECVDKCN	LLEGEPREFV	550
	ENSECIQCHP	ECLPQAMNIT	CTGRGPDNCI	QCAHYIDGPH	CVKTCPAGVM	600
	GENNTLVWKY	ADAGHVCHLC	HPNCTYGCTG	PGLEGCP TNG	PKIPSIATGM	650
	VGALLLLLVV	ALGIGLFMR	RHIVRKRTL	RLLQERELVE	PLTPSGEAPN	700
15	QALLRILKET	EFKKIKVLGS	GAFGTVYKGL	WIPEGEKVKI	PVAIKELREA	750
	TSPKANKEIL	DEAYVMASVD	NPHVCRLGI	CLTSTVQLIT	QLMPFGCLLD	800
	YVREHKDNIG	SQYLLNWCVQ	IAKGMNYLED	RRLVHRDLAA	RNVLVKTPQH	850
	VKITDFGLAK	LLGAEEKEYH	AEGGKVPIKW	MALESILHRI	YTHQSDVWSY	900
	GVTVWELMTF	GSKPYDGIPA	SEISSILEKG	ERLPQPPICT	IDVYMIMVKC	950
20	WMIDADSRPK	FRELIIEFSK	MARDPQRYLV	IQGDERMHL	SPTDSNFYRA	1000
	LMDEEDMDDV	VDADEYLIPQ	QGFSSPSTS	RTPLLSSLSA	TSNNSTVACI	1050
	DRNGLQSCPI	KEDSFLQRY	SDPTGALTED	SIDDTFLPVP	EYINQSVPKR	1100
	PAGSVQNPVY	HNQPLNPAPS	RDPHYQDPHS	TAVGNPEYLN	TVQPTCVNST	1150
	FDSPAHWQK	GSHQISLDNP	DYQQDFFPKE	AKPNGIFKGS	TAENAEYLRV	1200
25	APQSSEFIGA					1210

FIGURE 2

	<b>ATGCGACCCT</b>	CCGGGACGGC	CGGGGCAGCG	CTCCTGGCGC	TGCTGGCTGC	50
	GCTCTGCCCCG	GCGAGTCGGG	CTCTGGAGGA	AAAGAAAGTT	TGCCAAGGCA	100
	CGAGTAACAA	GCTCACGCAG	TTGGGCACTT	TTGAAGATCA	TTTTCTCAGC	150
5	CTCCAGAGGA	TGTTCAATAA	CTGTGAGGTG	GTCCTTGGGA	ATTTGGAAAT	200
	TACCTATGTG	CAGAGGAATT	ATGATCTTTC	CTTCTTAAAG	ACCATCCAGG	250
	AGGTGGCTGG	TTATGTCCTC	ATTGCCCTCA	ACACAGTGGA	GCGAATTCCT	300
	TTGGAAAACC	TGCAGATCAT	CAGAGGAAAT	ATGTACTACG	AAAATTCCTA	350
	TGCCTTAGCA	GTCTTATCTA	ACTATGATGC	AAATAAAACC	GGACTGAAGG	400
10	AGCTGCCCCAT	GAGAAATTTA	CAGGAAATCC	TGCATGGCGC	CGTGCGGTTC	450
	AGCAACAACC	CTGCCCTGTG	CAACGTGGAG	AGCATCCAGT	GGCGGGACAT	500
	AGTCAGCAGT	GACTTTCTCA	GCAACATGTC	GATGGACTTC	CAGAACCACC	550
	TGGGCAGCTG	CCAAAAGTGT	GATCCAAGCT	GTCCCAATGG	GAGCTGCTGG	600
	GGTGCAGGAG	AGGAGAACTG	CCAGAAACTG	ACCAAAATCA	TCTGTGCCCA	650
15	GCAGTGCTCC	GGGCGCTGCC	GTGGCAAGTC	CCCCAGTGAC	TGCTGCCACA	700
	ACCAGTGTGC	TGCAGGCTGC	ACAGGCCCCC	GGGAGAGCGA	CTGCCTGGTC	750
	TGCCGCAAAT	TCCGAGACGA	AGCCACGTGC	AAGGACACCT	GCCCCCCT	800
	CATGCTCTAC	AACCCACCA	CGTACCAGAT	GGATGTGAAC	CCCAGGGGCA	850
	AATACAGCTT	TGGTGCCACC	TGCGTGAAGA	AGTGTCCCCG	TAATTATGTG	900
20	GTGACAGATC	ACGGCTCGTG	CGTCCGAGCC	TGTGGGGCCG	ACAGCTATGA	950
	GATGGAGGAA	GACGGCGTCC	GCAAGTGTA	GAAGTGC	GAAAGTGGC	1000
	GCAAAGTGTG	TAACGGAATA	GGTATTGGTG	AATTTAAAGA	CTCACTCTCC	1050
	ATAAATGCTA	CGAATATTAA	ACACTTCAAA	AACTGCACCT	CCATCAGTGG	1100
	CGATCTCCAC	ATCCTGCCGG	TGGCATTTAG	GGGTGACTCC	TTCACACATA	1150
25	CTCCTCCTCT	GGATCCACAG	GAAGTGGATA	TTCTGAAAAC	CGTAAAGGAA	1200
	ATCACAGGGT	TTTTGCTGAT	TCAGGCTTGG	CCTGAAAACA	GGACGGACCT	1250
	CCATGCCTTT	GAGAACCTAG	AAATCATACG	CGGCAGGACC	AAGCAACATG	1300
	GTCAGTTTTT	TCTTGACGTC	GTCAGCCTGA	ACATAACATC	CTTGGGATTA	1350
30	CGCTCCCTCA	AGGAGATAAG	TGATGGAGAT	GTGATAATTT	CAGGAAACAA	1400

FIGURE 3A

	AAATTTGTGC	TATGCAAATA	CAATAAACTG	GAAAAAACTG	TTTGGGACCT	1450
	CCGGTCAGAA	AACCAAAATT	ATAAGCAACA	GAGGTGAAAA	CAGCTGCAAG	1500
	GCCACAGGCC	AGGTCTGCCA	TGCCTTGTGC	TCCCCCGAGG	GCTGCTGGGG	1550
5	CCCGGAGCCC	AGGGACTGCG	TCTCTTGCCG	GAATGTCAGC	CGAGGCAGGG	1600
	AATGCGTGGA	CAAGTGCAAG	CTTCTGGAGG	GTGAGCCAAG	GGAGTTTGTG	1650
	GAGAACTCTG	AGTGCATACA	GTGCCACCCA	GAGTGCCTGC	CTCAGGCCAT	1700
	GAACATCACC	TGCACAGGAC	GGGGACCAGA	CAACTGTATC	CAGTGTGCCC	1750
	ACTACATTGA	CGGCCCCCAC	TGCGTCAAGA	CCTGCCCCGGC	AGGAGTCATG	1800
10	GGAGAAAACA	ACACCCTGGT	CTGGAAGTAC	GCAGACGCCG	GCCATGTGTG	1850
	CCACCTGTGC	CATCCAAACT	GCACCTACGG	ATGCACTGGG	CCAGGTCTTG	1900
	AAGGCTGTCC	AACGAATGGG	CCTAAGATCC	CGTCCATCGC	CACTGGGATG	1950
	GTGGGGGCCC	TCCTCTTGCT	GCTGGTGGTG	GCCCTGGGGA	TCGGCCTCTT	2000
	CATGCGAAGG	CGCCACATCG	TTCGGAAGCG	CACGCTGCGG	AGGCTGCTGC	2050
15	AGGAGAGGGA	GCTTGTGGAG	CCTCTTACAC	CCAGTGGAGA	AGCTCCCAAC	2100
	CAAGCTCTCT	TGAGGATCTT	GAAGGAAACT	GAATTCAAAA	AGATCAAAGT	2150
	GCTGGGCTCC	GGTGCGTTTC	GCACGGTGTA	TAAGGGACTC	TGGATCCAG	2200
	AAGGTGAGAA	AGTTAAAATT	CCCGTCGCTA	TCAAGGAATT	AAGAGAAGCA	2250
	ACATCTCCGA	AAGCCAACAA	GGAAATCCTC	GATGAAGCCT	ACGTGATGGC	2300
20	CAGCGTGGAC	AACCCCCACG	TGTGCCGCCT	GCTGGGCATC	TGCCTCACCT	2350
	CCACCGTGCA	ACTCATCACG	CAGCTCATGC	CCTTCGGCTG	CCTCCTGGAC	2400
	TATGTCCGGG	AACACAAAGA	CAATATTGGC	TCCCAGTACC	TGCTCAACTG	2450
	GTGTGTGCAG	ATCGCAAAGG	GCATGAACTA	CTTGGAGGAC	CGTCGCTTGG	2500
	TGCACCGCGA	CCTGGCAGCC	AGGAACGTAC	TGGTGAAAAC	ACCGCAGCAT	2550
25	GTCAAGATCA	CAGATTTTGG	GCTGGCCAAA	CTGCTGGGTG	CGGAAGAGAA	2600
	AGAATACCAT	GCAGAAGGAG	GCAAAGTGCC	TATCAAGTGG	ATGGCATTGG	2650
	AATCAATTTT	ACACAGAATC	TATACCCACC	AGAGTGATGT	CTGGAGCTAC	2700
	GGGGTGACCG	TTTGGGAGTT	GATGACCTTT	GGATCCAAGC	CATATGACGG	2750
	AATCCCTGCC	AGCGAGATCT	CCTCCATCCT	GGAGAAAGGA	GAACGCCTCC	2800
30						

FIGURE 3B

	CTCAGCCACC	CATATGTACC	ATCGATGTCT	ACATGATCAT	GGTCAAGTGC	2850
	TGGATGATAG	ACGCAGATAG	TCGCCCCAAG	TTCCGTGAGT	TGATCATCGA	2900
	ATTCTCCAAA	ATGGCCCCGAG	ACCCCCAGCG	CTACCTTGTC	ATTCAGGGGG	2950
	ATGAAAGAAT	GCATTTGCCA	AGTCCTACAG	ACTCCAACCTT	CTACCGTGCC	3000
5	CTGATGGATG	AAGAAGACAT	GGACGACGTG	GTGGATGCCG	ACGAGTACCT	3050
	CATCCCACAG	CAGGGCTTCT	TCAGCAGCCC	CTCCACGTCA	CGGACTCCCC	3100
	TCCTGAGCTC	TCTGAGTGCA	ACCAGCAACA	ATTCCACCGT	GGCTTGCACTT	3150
	GATAGAAATG	GGCTGCAAAG	CTGTCCCATC	AAGGAAGACA	GCTTCTTGCA	3200
	GCGATACAGC	TCAGACCCCA	CAGGCGCCTT	GACTGAGGAC	AGCATAGACG	3250
10	ACACCTTCCT	CCCAGTGCCT	GAATACATAA	ACCAGTCCGT	TCCCCAAAAGG	3300
	CCCGCTGGCT	CTGTGCAGAA	TCCTGTCTAT	CACAATCAGC	CTCTGAACCC	3350
	CGCGCCCAGC	AGAGACCCAC	ACTACCAGGA	CCCCCACAGC	ACTGCAGTGG	3400
	GCAACCCCGA	GTATCTCAAC	ACTGTCCAGC	CCACCTGTGT	CAACAGCACA	3450
	TTCGACAGCC	CTGCCCCACTG	GGCCCAGAAA	GGCAGCCACC	AAATTAGCCT	3500
15	GGACAACCCT	GACTACCAGC	AGGACTTCTT	TCCCAAGGAA	GCCAAGCCAA	3550
	ATGGCATCTT	TAAGGGCTCC	ACAGCTGAAA	ATGCAGAATA	CCTAAGGGTC	3600
	GCGCCACAAA	GCAGTGAATT	TATTGGAGCA	<b>TGA</b>		3630

20

25

FIGURE 3C

### First PCR reaction

EGFR1+ EGFR3 primer

a.)  EGFR cDNA

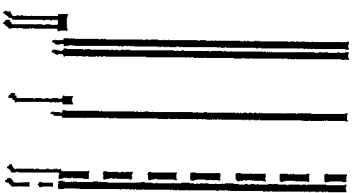
**PCR product a**  
encoding  
aa 1 to 24 fused  
to aa 313 to 319

EGFR2 + EGFR2220R primer

b.)  EGFR cDNA

**PCR product b**  
encoding  
aa 18 to 24 fused  
to aa 313 to 678

### Second PCR reaction

a.)  **PCR product a + b**  
**overlap extension**  
**mutant PCR product c**  
encoding aa 1 to 24 fused to aa 313 to 678  
with a stop codon in position 679


b.) EGFR1 + EGFR2220R primer  
 **amplification of PCR product c**

FIGURE 4

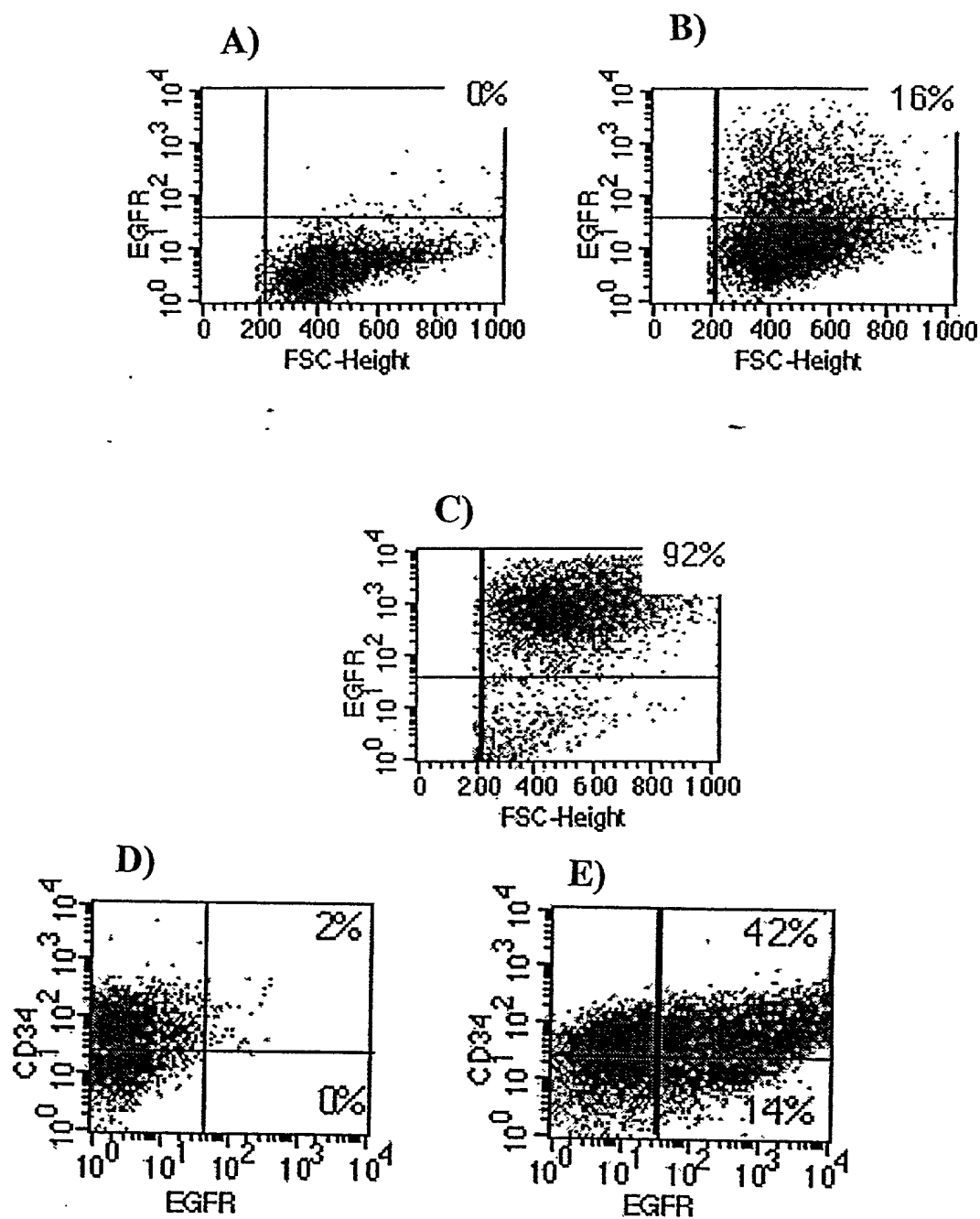


FIGURE 5

	MRELVNIPLV	HILTLVAFSG	TEKLPKAPVI	TTPLETVDAL	VEEVATFMCA	50
	VESYPQPEIS	WTRNKILIKL	FDTRYSIREN	GQLLTILSVE	DSDDGIYCCT	100
	ANNGVGGAVE	SCGALQVKMK	PKITRPPINV	KIIEGLKAVL	PCTTMGNPKP	150
5	SVSWIKGDSP	LRENSRIAVL	ESGSLRIHNV	QKEDAGQYRC	VAKNSLGTAY	200
	SKVVKLEFEV	FARILRAPES	HNVTFGSFVT	LHCTATGIPV	PTITWIENG	250
	AVSSGSIQES	VKDRVIDSRL	QLFITKPGLY	TCIATNKHGE	KFSTAKAAAT	300
	ISIAEWSKPQ	KDNKGYCAQY	RGEVCNAVLA	KDALVFLNTS	YADPEEAQEL	350
	LVHTAWNELK	VVSPVCRPAA	EALLCNHIFQ	ECSPGVVPTP	IPICREYCLA	400
10	VKELFCAKEW	LVMEETHRG	LYRSEMHLIS	VPKCSKLPSM	HWDPTACARL	450
	PHLDYNKENL	KTFPPMTSSK	PSVDIPNLPS	SSSSSFSVSP	TYSMTVIISI	500
	MSSFAIFVLL	TITTLYCCRR	RKQWKNKKRE	SAAVTLTTL	SELLLDRLHP	550
	NPMYQRMPLL	LNPKLISLEY	PRNNIEYVRD	IGEGAFGRVF	QARAPGLLPY	600
	EPFTMVAVKM	LKEEASADMQ	ADFQREAAAL	AEFDNPNIVK	LLGVCavgkp	650
15	MCLLFeyMay	GDLNEFLRSM	SPHTVCSLSH	SDLSMRAQVS	SPGPPPLSCA	700
	EQLCIARQVA	AGMAYLSERK	FVHRDLATRN	CLVGENMVVK	IADFGLSRNI	750
	YSADYYKANE	NDAIPiRWMP	PESIFYNRYT	TESDVWAYGV	VLWEIFSyGL	800
	QPYYGMAHEE	VIYYVRDgNI	LSCPENCpVE	LYNLMRLCWS	KLPADRPSFT	850
20	SIHRILERMC	ERAEGTVSV				869

FIGURE 6



DECLARATION AND POWER OF ATTORNEY FOR U.S. PATENT APPLICATIONS

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name,  
and

I believe I am an original, first and joint inventor of the subject matter which is claimed  
and for which a patent is sought on the invention entitled

Selectable Cell Surface Marker Genes

the specification of which is attached hereto.

I hereby state that I have reviewed and understand the contents of the above identified  
specification, including the claims.

I acknowledge my duty to disclose all information which is known by me to be material to  
the patentability of this application as defined in 37 C.F.R. §1.56.

I hereby claim the benefit under 35 U.S.C. §119(a)-(d) or §365(b) of any foreign  
application(s) for patent or inventor's certificate listed below and under 35 U.S.C. §365(a) of any  
PCT international application(s) designating at least one country other than the United States  
listed below and have also listed below any foreign application(s) for patent or inventor's  
certificate or any PCT international application(s) designating at least one country other than the  
United States for the same subject matter and having a filing date before that of the application  
the priority of which is claimed for that subject matter:

None

I hereby claim the benefit under 35 USC §119(e) of any United States provisional  
application(s) identified below:

<u>Application No.</u>	<u>Filing Date</u>
1. Provisional Application No. 60/166,594 For "Selectable Cell Surface Marker Genes	November 19, 1999
2. Provisional Application No. to be assigned (filed as Application No. 09/444,038 for "Selective Marker Genes," and subject	November 19, 1999

to a Petition for Conversion to Provisional Application, filed November 16, 2000)

3. Provisional Application No. to be assigned (filed as Application No. 09/539,248 for "Selectable Marker Genes," and subject to a Petition for Conversion to Provisional Application, filed November 16, 2000).

March 30, 2000

I hereby claim the benefit under 35 U.S.C. §120 of any United States application(s) listed below and under 35 U.S.C. §365(c) of any PCT international application(s) designating the United States listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in said prior application(s) in the manner required by the first paragraph of 35 U.S.C. §112, I acknowledge the duty to disclose all information known by me to be material to patentability as defined in 37 C.F.R. §1.56 which became available between the filing date(s) of the prior application(s) and the national or PCT international filing date of this application:

None

I hereby appoint the attorneys and agents associated with Customer No. 001095, respectively and individually, as my attorneys and agents, with full power of substitution and revocation, to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith.

Please address all communications to the address associated with Customer No. 001095, which is currently Thomas Hoxie, Novartis Corporation, Patent and Trademark Dept., 564 Morris Avenue, Summit, NJ 07901-1027.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under 18 U.S.C. §1001 and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

FIRST JOINT INVENTOR:

Full name : **Susanne Dagmar Pippig**

Signature : \_\_\_\_\_

Date : \_\_\_\_\_  
(MM/DD/YY)

Citizenship : Germany

Residence : San Francisco, California

P.O. Address : 1495 7th Avenue  
Apartment 2  
San Francisco, California 94122

SECOND JOINT INVENTOR:

Full name : **Gabor Veres**

Signature : \_\_\_\_\_

Date : \_\_\_\_\_  
(MM/DD/YY)

Citizenship : Hungary

Residence : Palo Alto, California

P.O. Address : 112 Walter Hays Drive  
Palo Alto, California 94303

**IMPORTANT:** Before this declaration is signed, the patent application (the specification, the claims and this declaration) must be read and understood by each person signing it, and no changes may be made in the application after this declaration has been signed.

31192prj.ST25.txt  
SEQUENCE LISTING



<110> NOVARTIS AG  
VERES, GABOR  
PIPPIG, SUSANNE

<120> selectable cell surface marker genes

<130> 4-31192

<150> us 60/166594

<151> 1999-11-19

<150> us 09/539248

<151> 2000-03-30

<160> 16

<170> PatentIn version 3.0

<210> 1

<211> 3633

<212> DNA

<213> EGFR

<400> 1

atgcgaccct cggggacggc cggggcagcg ctcttggcgc tgctggctgc gctctgcccc  
60

gcgagtcggg ctctggagga aaagaaagtt tgccaaggca cgagtaacaa gctcacgcag 1  
20

ttgggcactt ttgaagatca ttttctcagc ctccagagga tgttcaataa ctgtgagggtg 1  
80

gtccttggga atttggaat tacctatgtg cagaggaatt atgatcttct cttcttaaag 2  
40

accatccagg aggtggctgg ttatgtcctc attgccctca acacagtgga gcgaattcct 3  
00

ttggaaaacc tgcagatcat cagaggaaat atgtactacg aaaattccta tgccttagca 3  
60

gtcttatcta actatgatgc aaataaaacc ggactgaagg agctgcccat gagaaattta 4  
20

caggaaatcc tgcattggcg cgtgcgggtc agcaacaacc ctgccctgtg caacgtggag 4

## 31192prj.ST25.txt

80

agcatccagt ggcgggacat agtcagcagt gactttctca gcaacatgtc gatggacttc 5  
40

cagaaccacc tgggcagctg ccaaaagtgt gatccaagct gtcccaatgg gagctgctgg 6  
00

ggtgcaggag aggagaactg ccagaaaactg accaaaatca tctgtgccca gcagtgctcc 6  
60

gggcgctgcc gtggcaagtc cccagtgac tgctgccaca accagtgtgc tgcaggctgc 7  
20

acaggccccc gggagagcga ctgcctggtc tgccgcaaat tccgagacga agccacgtgc 7  
80

aaggacacct gccccccact catgctctac aacccaccca cgtaccagat ggatgtgaac 8  
40

cccgagggca aatacagctt tgggtgccacc tgcgtgaaga agtgtccccg taattatgtg 9  
00

gtgacagatc acggctcgtg cgtccgagcc tgtggggccg acagctatga gatggaggaa 9  
60

gacggcgtcc gcaagtgtaa gaagtgcgaa gggccttgcc gcaaagtgtg taacggaata 10  
20

ggtattggtg aatttaaaga ctactctcc ataaatgcta cgaatattaa acacttcaaa 10  
80

aactgcacct ccatcagtgg cgatctccac atcctgccgg tggcatttag gggtgactcc 11  
40

ttcacacata ctctcctct ggatccacag gaactggata ttctgaaaac cgtaaaggaa 12  
00

atcacagggg ttttgctgat tcaggcttgg cctgaaaaca ggacggacct ccatgccttt 12  
60

gagaacctag aaatcatacg cggcaggacc aagcaacatg gtcagttttc tcttgcagtc 13  
20

gtcagcctga acataacatc cttgggatta cgctccctca aggagataag tgatggagat 13  
80

gtgataattt caggaaacaa aaatttgtgc tatgcaaata caataaactg gaaaaaactg 14

## 31192prj.ST25.txt

40

tttgggacct ccggtcagaa aacccaaatt ataagcaaca gaggtgaaaa cagctgcaag 15  
00

gccacaggcc aggtctgcca tgccttgtgc tccccgagg gctgctgggg cccggagccc 15  
60

agggactgcg tctcttgccg gaatgtcagc cgaggcaggg aatgcgtgga caagtgcaag 16  
20

cttctggagg gtgagccaag ggagtttgtg gagaactctg agtgcataca gtgccacca 16  
80

gagtgcctgc ctcaggccat gaacatcacc tgcacaggac ggggaccaga caactgtatc 17  
40

cagtgtgccc actacattga cggccccac tgcgtcaaga cctgcccggc aggagtcattg 18  
00

ggagaaaaca acaccctggc ctggaagtac gcagacgccg gccatgtgtg ccacctgtgc 18  
60

catccaaact gcacctacgg atgcactggg ccaggctctg aaggctgtcc aacgaatggg 19  
20

cctaagatcc cgtccatcgc cactgggatg gtggggggccc tctcttgct gctgggtggtg 19  
80

gccctgggga tcggcctctt catgcgaagg cgccacatcg ttcggaagcg cacgctgcgg 20  
40

aggctgctgc aggagagggg gcttgtggag cctcttacac ccagtggaga agctcccaac 21  
00

caagctctct tgaggatctt gaaggaaact gaattcaaaa agatcaaagt gctgggctcc 21  
60

ggtgcgttcg gcacggtgta taagggactc tggatcccag aaggtagaaa agttaaaatt 22  
20

cccgctgcta tcaaggaatt aagagaagca acatctccga aagccaacaa ggaaatcctc 22  
80

gatgaagcct acgtgatggc cagcgtggac aacccccacg tgtgccgcct gctgggcatc 23  
40

tgcctcacct ccaccgtgca actcatcacg cagctcatgc ccttcggctg cctcctggac 24

31192prj.ST25.txt

00

tatgtccggg aacacaaaga caatattggc tcccagtacc tgctcaactg gtgtgtgcag 24  
60

atcgcaaagg gcatgaacta cttggaggac cgtcgcttgg tgcaccgcga cctggcagcc 25  
20

aggaacgtac tggtgaaaac accgcagcat gtcaagatca cagatttttg gctggccaaa 25  
80

ctgctgggtg cggaagagaa agaataccat gcagaaggag gcaaagtgcc tatcaagtgg 26  
40

atggcattgg aatcaatttt acacagaatc tatacccacc agagtgatgt ctggagctac 27  
00

ggggtgaccg tttgggagtt gatgaccttt ggatccaagc catatgacgg aatccctgcc 27  
60

agcgagatct cctccatcct ggagaaagga gaacgcctcc ctcagccacc catatgtacc 28  
20

atcgatgtct acatgatcat ggtcaagtgc tggatgatag acgcagatag tcgccccaaag 28  
80

ttccgtgagt tgatcatcga attctccaaa atggccccgag acccccagcg ctaccttgtc 29  
40

attcaggggg atgaaagaat gcatttgcca agtcctacag actccaactt ctaccgtgcc 30  
00

ctgatggatg aagaagacat ggacgacgtg gtggatgccg acgagtacct catcccacag 30  
60

cagggttctt tcagcagccc ctccacgtca cggactcccc tcctgagctc tctgagtgca 31  
20

accagcaaca attccaccgt ggcttgcatt gatagaaatg ggctgcaaag ctgtcccatc 31  
80

aaggaagaca gcttcttgca gcgatacagc tcagacccca caggcgcctt gactgaggac 32  
40

agcatagacg acaccttcct cccagtgcct gaatacataa accagtccgt tccccaaagg 33  
00

cccgtggct ctgtgcagaa tcctgtctat cacaatcagc ctctgaacct cgcgcccagc 33

60

agagacccac actaccagga cccccacagc actgcagtgg gcaaccccga gtatctcaac 34  
20

actgtccagc ccacctgtgt caacagcaca ttgcacagcc ctgcccactg ggcccagaaa 34  
80

ggcagccacc aaattagcct ggacaaccct gactaccagc aggacttctt tcccaaggaa 35  
40

gccaagccaa atggcatctt taagggctcc acagctgaaa atgcagaata cctaagggtc 36  
00

gcgccacaaa gcagtgaatt tattggagca tga 36  
33

<210> 2  
<211> 1210  
<212> PRT  
<213> EGFR

<400> 2

Met Arg Pro Ser Gly Thr Ala Gly Ala Ala Leu Leu Ala Leu Leu Ala  
1 5 10 15

Ala Leu Cys Pro Ala Ser Arg Ala Leu Glu Glu Lys Lys Val Cys Gln  
20 25 30

Gly Thr Ser Asn Lys Leu Thr Gln Leu Gly Thr Phe Glu Asp His Phe  
35 40 45

Leu Ser Leu Gln Arg Met Phe Asn Asn Cys Glu Val Val Leu Gly Asn  
50 55 60

Leu Glu Ile Thr Tyr Val Gln Arg Asn Tyr Asp Leu Ser Phe Leu Lys  
65 70 75 80

Thr Ile Gln Glu Val Ala Gly Tyr Val Leu Ile Ala Leu Asn Thr Val  
85 90 95

Glu Arg Ile Pro Leu Glu Asn Leu Gln Ile Ile Arg Gly Asn Met Tyr  
100 105 110

Tyr Glu Asn Ser Tyr Ala Leu Ala Val Leu Ser Asn Tyr Asp Ala Asn  
115 120 125



31192prj.ST25.txt

Lys	Thr	Gly	Leu	Lys	Glu	Leu	Pro	Met	Arg	Asn	Leu	Gln	Glu	Ile	Leu
	130					135					140				
His	Gly	Ala	Val	Arg	Phe	Ser	Asn	Asn	Pro	Ala	Leu	Cys	Asn	Val	Glu
145					150					155					160
Ser	Ile	Gln	Trp	Arg	Asp	Ile	Val	Ser	Ser	Asp	Phe	Leu	Ser	Asn	Met
				165					170					175	
Ser	Met	Asp	Phe	Gln	Asn	His	Leu	Gly	Ser	Cys	Gln	Lys	Cys	Asp	Pro
			180					185					190		
Ser	Cys	Pro	Asn	Gly	Ser	Cys	Trp	Gly	Ala	Gly	Glu	Glu	Asn	Cys	Gln
		195					200					205			
Lys	Leu	Thr	Lys	Ile	Ile	Cys	Ala	Gln	Gln	Cys	Ser	Gly	Arg	Cys	Arg
	210					215						220			
Gly	Lys	Ser	Pro	Ser	Asp	Cys	Cys	His	Asn	Gln	Cys	Ala	Ala	Gly	Cys
225					230					235					240
Thr	Gly	Pro	Arg	Glu	Ser	Asp	Cys	Leu	Val	Cys	Arg	Lys	Phe	Arg	Asp
				245					250					255	
Glu	Ala	Thr	Cys	Lys	Asp	Thr	Cys	Pro	Pro	Leu	Met	Leu	Tyr	Asn	Pro
			260					265					270		
Thr	Thr	Tyr	Gln	Met	Asp	Val	Asn	Pro	Glu	Gly	Lys	Tyr	Ser	Phe	Gly
		275					280					285			
Ala	Thr	Cys	Val	Lys	Lys	Cys	Pro	Arg	Asn	Tyr	Val	Val	Thr	Asp	His
	290					295					300				
Gly	Ser	Cys	Val	Arg	Ala	Cys	Gly	Ala	Asp	Ser	Tyr	Glu	Met	Glu	Glu
305					310					315					320
Asp	Gly	Val	Arg	Lys	Cys	Lys	Lys	Cys	Glu	Gly	Pro	Cys	Arg	Lys	Val
				325					330					335	
Cys	Asn	Gly	Ile	Gly	Ile	Gly	Glu	Phe	Lys	Asp	Ser	Leu	Ser	Ile	Asn
			340					345					350		
Ala	Thr	Asn	Ile	Lys	His	Phe	Lys	Asn	Cys	Thr	Ser	Ile	Ser	Gly	Asp
		355					360					365			
Leu	His	Ile	Leu	Pro	Val	Ala	Phe	Arg	Gly	Asp	Ser	Phe	Thr	His	Thr
	370					375					380				

31192prj.ST25.txt

Pro	Pro	Leu	Asp	Pro	Gln	Glu	Leu	Asp	Ile	Leu	Lys	Thr	Val	Lys	Glu	385	390	395	400
Ile	Thr	Gly	Phe	Leu	Leu	Ile	Gln	Ala	Trp	Pro	Glu	Asn	Arg	Thr	Asp	405	410	415	
Leu	His	Ala	Phe	Glu	Asn	Leu	Glu	Ile	Ile	Arg	Gly	Arg	Thr	Lys	Gln	420	425	430	
His	Gly	Gln	Phe	Ser	Leu	Ala	Val	Val	Ser	Leu	Asn	Ile	Thr	Ser	Leu	435	440	445	
Gly	Leu	Arg	Ser	Leu	Lys	Glu	Ile	Ser	Asp	Gly	Asp	Val	Ile	Ile	Ser	450	455	460	
Gly	Asn	Lys	Asn	Leu	Cys	Tyr	Ala	Asn	Thr	Ile	Asn	Trp	Lys	Lys	Leu	465	470	475	480
Phe	Gly	Thr	Ser	Gly	Gln	Lys	Thr	Lys	Ile	Ile	Ser	Asn	Arg	Gly	Glu	485	490	495	
Asn	Ser	Cys	Lys	Ala	Thr	Gly	Gln	Val	Cys	His	Ala	Leu	Cys	Ser	Pro	500	505	510	
Glu	Gly	Cys	Trp	Gly	Pro	Glu	Pro	Arg	Asp	Cys	Val	Ser	Cys	Arg	Asn	515	520	525	
Val	Ser	Arg	Gly	Arg	Glu	Cys	Val	Asp	Lys	Cys	Asn	Leu	Leu	Glu	Gly	530	535	540	
Glu	Pro	Arg	Glu	Phe	Val	Glu	Asn	Ser	Glu	Cys	Ile	Gln	Cys	His	Pro	545	550	555	560
Glu	Cys	Leu	Pro	Gln	Ala	Met	Asn	Ile	Thr	Cys	Thr	Gly	Arg	Gly	Pro	565	570	575	
Asp	Asn	Cys	Ile	Gln	Cys	Ala	His	Tyr	Ile	Asp	Gly	Pro	His	Cys	Val	580	585	590	
Lys	Thr	Cys	Pro	Ala	Gly	Val	Met	Gly	Glu	Asn	Asn	Thr	Leu	Val	Trp	595	600	605	
Lys	Tyr	Ala	Asp	Ala	Gly	His	Val	Cys	His	Leu	Cys	His	Pro	Asn	Cys	610	615	620	
Thr	Tyr	Gly	Cys	Thr	Gly	Pro	Gly	Leu	Glu	Gly	Cys	Pro	Thr	Asn	Gly	625	630	635	640

## 31192prj.ST25.txt

Pro Lys Ile Pro Ser Ile Ala Thr Gly Met Val Gly Ala Leu Leu Leu  
 645 650 655  
 Leu Leu Val Val Ala Leu Gly Ile Gly Leu Phe Met Arg Arg Arg His  
 660 665 670  
 Ile Val Arg Lys Arg Thr Leu Arg Arg Leu Leu Gln Glu Arg Glu Leu  
 675 680 685  
 Val Glu Pro Leu Thr Pro Ser Gly Glu Ala Pro Asn Gln Ala Leu Leu  
 690 695 700  
 Arg Ile Leu Lys Glu Thr Glu Phe Lys Lys Ile Lys Val Leu Gly Ser  
 705 710 715 720  
 Gly Ala Phe Gly Thr Val Tyr Lys Gly Leu Trp Ile Pro Glu Gly Glu  
 725 730 735  
 Lys Val Lys Ile Pro Val Ala Ile Lys Glu Leu Arg Glu Ala Thr Ser  
 740 745 750  
 Pro Lys Ala Asn Lys Glu Ile Leu Asp Glu Ala Tyr Val Met Ala Ser  
 755 760 765  
 Val Asp Asn Pro His Val Cys Arg Leu Leu Gly Ile Cys Leu Thr Ser  
 770 775 780  
 Thr Val Gln Leu Ile Thr Gln Leu Met Pro Phe Gly Cys Leu Leu Asp  
 785 790 795 800  
 Tyr Val Arg Glu His Lys Asp Asn Ile Gly Ser Gln Tyr Leu Leu Asn  
 805 810 815  
 Trp Cys Val Gln Ile Ala Lys Gly Met Asn Tyr Leu Glu Asp Arg Arg  
 820 825 830  
 Leu Val His Arg Asp Leu Ala Ala Arg Asn Val Leu Val Lys Thr Pro  
 835 840 845  
 Gln His Val Lys Ile Thr Asp Phe Gly Leu Ala Lys Leu Leu Gly Ala  
 850 855 860  
 Glu Glu Lys Glu Tyr His Ala Glu Gly Gly Lys Val Pro Ile Lys Trp  
 865 870 875 880  
 Met Ala Leu Glu Ser Ile Leu His Arg Ile Tyr Thr His Gln Ser Asp  
 885 890 895

Page 9

31192prj.ST25.txt

Pro Thr Cys Val Asn Ser Thr Phe Asp Ser Pro Ala His Trp Ala  
1145 1150 1155

Gln Lys Gly Ser His Gln Ile Ser Leu Asp Asn Pro Asp Tyr Gln  
1160 1165 1170

Gln Asp Phe Phe Pro Lys Glu Ala Lys Pro Asn Gly Ile Phe Lys  
1175 1180 1185

Gly Ser Thr Ala Glu Asn Ala Glu Tyr Leu Arg Val Ala Pro Gln  
1190 1195 1200

Ser Ser Glu Phe Ile Gly Ala  
1205 1210

<210> 3  
<211> 31  
<212> DNA  
<213> primer

<400> 3  
ctaggctagc atgcgaccct ccgggacggc c  
31

<210> 4  
<211> 42  
<212> DNA  
<213> primer

<400> 4  
ctctgcccgg cgactcgggc tgacagctat gagatggagg aa  
42

<210> 5  
<211> 39  
<212> DNA  
<213> primer

<400> 5  
ttcctccatc tcatagctag ccgactcgc cgggcagag  
39

<210> 6  
<211> 31

31192prj.ST25.txt

<212> DNA  
<213> primer

<400> 6  
ggatataccta cgtgcgcttc cgaacgatgt g  
31

<210> 7  
<211> 2607  
<212> DNA  
<213> MuSK

<400> 7  
atgagagagc tcgtcaacat tccactggta catattctta ctctgggtgc cttcagcgga  
60

actgagaaac ttccaaaagc tctgtgcatc accactcctc ttgaaacagt ggatgcctta 1  
20

gttgaagaag tggctacttt catgtgtgca gtggaatcct acccccagcc tgagatttcc 1  
80

tggactagaa ataaaattct cattaaactc tttgacaccc ggtacagcat ccgggagaat 2  
40

gggcagctcc tcaccatcct gagtgtggaa gacagtgatg atggcattta ctgctgcacg 3  
00

gccaacaatg gtgtgggagg agctgtggag agttgtggag ccctgcaagt gaagatgaaa 3  
60

cctaaaataa ctgctcctcc cataaatgtg aaaataatag agggattaaa agcagtccta 4  
20

ccatgtacta caatgggtaa tcccaaacca tcagtgtcctt ggataaaggg agacagccct 4  
80

ctcagggaaa attcccgaat tgcagttcctt gaatctggga gcttgaggat tcataacgta 5  
40

caaaaggaag atgcaggaca gtatcgatgt gtggcaaaaa acagcctcgg gacagcatat 6  
00

tccaaagtgg tgaagctgga atttgaggtt tttgccagga tcttgccgggc tcttgaatcc 6  
60

cacaatgtca cctttggctc ctttgtgacc ctgcaactgta cagcaacagg cattcctgtc 7

## 31192prj.ST25.txt

20

cccaccatca cctggattga aaacggaaat gctgttttctt ctgggtccat tcaagagagt 7  
80

gtgaaagacc gagtgattga ctcaagactg cagctgttta tcaccaagcc aggactctac 8  
40

acatgcatag ctaccaataa gcatggggag aagttcagta ctgccaaggc tgcagccacc 9  
00

atcagcatag cagaatggag taaaccacag aaagataaca aaggctactg cgcccagtac 9  
60

agaggggagg tgtgtaatgc agtcctggca aaagatgctc ttgtttttct caacacctcc 10  
20

tatgctggacc ctgaggaggc ccaagagcta ctgggtccaca cggcctggaa tgaactgaaa 10  
80

gtagtgagcc cagtctgccg gccagctgct gaggctttgt tgtgtaacca catcttccag 11  
40

gagtgcagtc ctggagtagt gcctactcct attcccattt gcagagagta ctgcttggca 12  
00

gtaaaggagc tcttctgctc aaaagaatgg ctggtaatgg aagagaagac ccacagagga 12  
60

ctctacagat ccgagatgca tttgctgtcc gtgccaaaat gcagcaagct tcccagcatg 13  
20

cattgggacc ccacggcctg tgccagactg ccacatctag attataacaa agaaaaccta 13  
80

aaaacattcc caccaatgac gtcctcaaag ccaagtgtgg acattccaaa tctgccttcc 14  
40

tcctcctctt cttccttctc tgtctcacct acatactcca tgactgtaat aatctccatc 15  
00

atgtccagct ttgcaatatt tgtgcttctt accataacta ctctctattg ctgccgaaga 15  
60

agaaaacaat ggaaaaataa gaaaagagaa tcagcagcag taaccctcac cacactgcct 16  
20

tctgagctct tactagatag acttcatccc aaccccatgt accagaggat gccgctcctt 16

31192prj.ST25.txt

80

ctgaacccca aattgctcag cctggagtat ccaaggaata acattgaata tgtgagagac 17  
40

atcggagagg gagcgtttgg aagggtgttt caagcaaggg caccaggctt acttccctat 18  
00

gaacctttca ctatggtggc agtaaagatg ctcaaagaag aagcctcggc agatatgcaa 18  
60

gcggactttc agagggaggc agccctcatg gcagaatttg acaaccctaa cattgtgaag 19  
20

ctattaggag tgtgtgctgt cgggaagcca atgtgcctgc tctttgaata catggcctat 19  
80

ggtgacctca atgagttcct ccgcagcatg tcccctcaca ccgtgtgcag cctcagtcac 20  
40

agtgacttgt ctatgagggc tcaggtctcc agccctgggc cccaccacct ctctgtgct 21  
00

gagcagcttt gcattgccag gcaggtggca gctggcatgg cttacctctc agaacgtaag 21  
60

tttgttcacc gagatttagc caccaggaac tgccctggtgg gcgagaacat ggtggtgaaa 22  
20

attgccgact ttggcctctc caggaaacac tactcagcag actactacaa agctaataaa 22  
80

aacgacgcta tccctatccg ttggatgcc aagagatcca ttttttataa ccgctacact 23  
40

acagagtctg atgtgtgggc ctatggcgtg gtcctctggg agatcttctc ctatggcctg 24  
00

cagccctact atgggatggc ccatgaggag gtcatttact acgtgcgaga tggcaacatc 24  
60

ctctcctgcc ctgagaactg ccccgaggag ctgtacaatc tcatgcgtct atgttggagc 25  
20

aagctgcctg cagacagacc cagtttcacc agtattcacc gaattctgga acgcatgtgt 25  
80

gagagggcag agggaaactgt gactgtgc 26



07

<210> 8  
 <211> 869  
 <212> PRT  
 <213> MuSK

<400> 8

Met	Arg	Glu	Leu	Val	Asn	Ile	Pro	Leu	Val	His	Ile	Leu	Thr	Leu	Val	1	5	10	15
Ala	Phe	Ser	Gly	Thr	Glu	Lys	Leu	Pro	Lys	Ala	Pro	Val	Ile	Thr	Thr	20	25	30	
Pro	Leu	Glu	Thr	Val	Asp	Ala	Leu	Val	Glu	Glu	Val	Ala	Thr	Phe	Met	35	40	45	
Cys	Ala	Val	Glu	Ser	Tyr	Pro	Gln	Pro	Glu	Ile	Ser	Trp	Thr	Arg	Asn	50	55	60	
Lys	Ile	Leu	Ile	Lys	Leu	Phe	Asp	Thr	Arg	Tyr	Ser	Ile	Arg	Glu	Asn	65	70	75	80
Gly	Gln	Leu	Leu	Thr	Ile	Leu	Ser	Val	Glu	Asp	Ser	Asp	Asp	Gly	Ile	85	90	95	
Tyr	Cys	Cys	Thr	Ala	Asn	Asn	Gly	Val	Gly	Gly	Ala	Val	Glu	Ser	Cys	100	105	110	
Gly	Ala	Leu	Gln	Val	Lys	Met	Lys	Pro	Lys	Ile	Thr	Arg	Pro	Pro	Ile	115	120	125	
Asn	Val	Lys	Ile	Ile	Glu	Gly	Leu	Lys	Ala	Val	Leu	Pro	Cys	Thr	Thr	130	135	140	
Met	Gly	Asn	Pro	Lys	Pro	Ser	Val	Ser	Trp	Ile	Lys	Gly	Asp	Ser	Pro	145	150	155	160
Leu	Arg	Glu	Asn	Ser	Arg	Ile	Ala	Val	Leu	Glu	Ser	Gly	Ser	Leu	Arg	165	170	175	
Ile	His	Asn	Val	Gln	Lys	Glu	Asp	Ala	Gly	Gln	Tyr	Arg	Cys	Val	Ala	180	185	190	
Lys	Asn	Ser	Leu	Gly	Thr	Ala	Tyr	Ser	Lys	Val	Val	Lys	Leu	Glu	Phe	195	200	205	

## 31192prj.ST25.txt

Glu Val Phe Ala Arg Ile Leu Arg Ala Pro Glu Ser His Asn Val Thr  
 210 215 220  
 Phe Gly Ser Phe Val Thr Leu His Cys Thr Ala Thr Gly Ile Pro Val  
 225 230 235 240  
 Pro Thr Ile Thr Trp Ile Glu Asn Gly Asn Ala Val Ser Ser Gly Ser  
 245 250 255  
 Ile Gln Glu Ser Val Lys Asp Arg Val Ile Asp Ser Arg Leu Gln Leu  
 260 265 270  
 Phe Ile Thr Lys Pro Gly Leu Tyr Thr Cys Ile Ala Thr Asn Lys His  
 275 280 285  
 Gly Glu Lys Phe Ser Thr Ala Lys Ala Ala Thr Ile Ser Ile Ala  
 290 295 300  
 Glu Trp Ser Lys Pro Gln Lys Asp Asn Lys Gly Tyr Cys Ala Gln Tyr  
 305 310 315 320  
 Arg Gly Glu Val Cys Asn Ala Val Leu Ala Lys Asp Ala Leu Val Phe  
 325 330 335  
 Leu Asn Thr Ser Tyr Ala Asp Pro Glu Glu Ala Gln Glu Leu Leu Val  
 340 345 350  
 His Thr Ala Trp Asn Glu Leu Lys Val Val Ser Pro Val Cys Arg Pro  
 355 360 365  
 Ala Ala Glu Ala Leu Leu Cys Asn His Ile Phe Gln Glu Cys Ser Pro  
 370 375 380  
 Gly Val Val Pro Thr Pro Ile Pro Ile Cys Arg Glu Tyr Cys Leu Ala  
 385 390 395 400  
 Val Lys Glu Leu Phe Cys Ala Lys Glu Trp Leu Val Met Glu Glu Lys  
 405 410 415  
 Thr His Arg Gly Leu Tyr Arg Ser Glu Met His Leu Leu Ser Val Pro  
 420 425 430  
 Lys Cys Ser Lys Leu Pro Ser Met His Trp Asp Pro Thr Ala Cys Ala  
 435 440 445  
 Arg Leu Pro His Leu Asp Tyr Asn Lys Glu Asn Leu Lys Thr Phe Pro  
 450 455 460

## 31192prj.ST25.txt

Pro Met Thr Ser Ser Lys Pro Ser Val Asp Ile Pro Asn Leu Pro Ser  
 465 470 475 480  
 Ser Ser Ser Ser Ser Phe Ser Val Ser Pro Thr Tyr Ser Met Thr Val  
 485 490 495  
 Ile Ile Ser Ile Met Ser Ser Phe Ala Ile Phe Val Leu Leu Thr Ile  
 500 505 510  
 Thr Thr Leu Tyr Cys Cys Arg Arg Arg Lys Gln Trp Lys Asn Lys Lys  
 515 520 525  
 Arg Glu Ser Ala Ala Val Thr Leu Thr Thr Leu Pro Ser Glu Leu Leu  
 530 535 540  
 Leu Asp Arg Leu His Pro Asn Pro Met Tyr Gln Arg Met Pro Leu Leu  
 545 550 555 560  
 Leu Asn Pro Lys Leu Leu Ser Leu Glu Tyr Pro Arg Asn Asn Ile Glu  
 565 570 575  
 Tyr Val Arg Asp Ile Gly Glu Gly Ala Phe Gly Arg Val Phe Gln Ala  
 580 585 590  
 Arg Ala Pro Gly Leu Leu Pro Tyr Glu Pro Phe Thr Met Val Ala Val  
 595 600 605  
 Lys Met Leu Lys Glu Glu Ala Ser Ala Asp Met Gln Ala Asp Phe Gln  
 610 615 620  
 Arg Glu Ala Ala Leu Met Ala Glu Phe Asp Asn Pro Asn Ile Val Lys  
 625 630 635 640  
 Leu Leu Gly Val Cys Ala Val Gly Lys Pro Met Cys Leu Leu Phe Glu  
 645 650 655  
 Tyr Met Ala Tyr Gly Asp Leu Asn Glu Phe Leu Arg Ser Met Ser Pro  
 660 665 670  
 His Thr Val Cys Ser Leu Ser His Ser Asp Leu Ser Met Arg Ala Gln  
 675 680 685  
 Val Ser Ser Pro Gly Pro Pro Pro Leu Ser Cys Ala Glu Gln Leu Cys  
 690 695 700  
 Ile Ala Arg Gln Val Ala Ala Gly Met Ala Tyr Leu Ser Glu Arg Lys  
 705 710 715 720

31192prj.ST25.txt

Phe Val His Arg Asp Leu Ala Thr Arg Asn Cys Leu Val Gly Glu Asn  
725 730 735

Met Val Val Lys Ile Ala Asp Phe Gly Leu Ser Arg Asn Ile Tyr Ser  
740 745 750

Ala Asp Tyr Tyr Lys Ala Asn Glu Asn Asp Ala Ile Pro Ile Arg Trp  
755 760 765

Met Pro Pro Glu Ser Ile Phe Tyr Asn Arg Tyr Thr Thr Glu Ser Asp  
770 775 780

Val Trp Ala Tyr Gly Val Val Leu Trp Glu Ile Phe Ser Tyr Gly Leu  
785 790 795 800

Gln Pro Tyr Tyr Gly Met Ala His Glu Glu Val Ile Tyr Tyr Val Arg  
805 810 815

Asp Gly Asn Ile Leu Ser Cys Pro Glu Asn Cys Pro Val Glu Leu Tyr  
820 825 830

Asn Leu Met Arg Leu Cys Trp Ser Lys Leu Pro Ala Asp Arg Pro Ser  
835 840 845

Phe Thr Ser Ile His Arg Ile Leu Glu Arg Met Cys Glu Arg Ala Glu  
850 855 860

Gly Thr Val Ser Val  
865

<210> 9  
<211> 25  
<212> DNA  
<213> primer

<400> 9  
cgtcctgcgt gagcctggat taatc  
25

<210> 10  
<211> 25  
<212> DNA  
<213> primer

<400> 10  
gcctggatta atcatgagag agctc

25

<210> 11  
<211> 42  
<212> DNA  
<213> primer

<400> 11  
cgaggcctgt cttcaacctt agacactcac agttccctct gc  
42

<210> 12  
<211> 27  
<212> DNA  
<213> primer

<400> 12  
cggcctgtgc cagactgccca catctag  
27

<210> 13  
<211> 32  
<212> DNA  
<213> primer

<400> 13  
cgtctagggtg agggttactg ctgctgattc tc  
32

<210> 14  
<211> 38  
<212> DNA  
<213> primer

<400> 14  
ggttaaccct attcaatggt attccttgga tactccag  
38

<210> 15  
<211> 24  
<212> DNA  
<213> primer

31192prj.ST25.txt

<400> 15  
cttccaaaag ctcctgtcat cacc  
24

<210> 16  
<211> 26  
<212> DNA  
<213> primer

<400> 16  
ccagtcatgg agtatgtagg tgagac  
26

Docket Number 4-31192/CIP	
FILING BY "EXPRESS MAIL" UNDER 37 CFR 1.10	
EL489123769US Express Mail Label Number	November 17, 2000 Date of Deposit

## IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

IN RE APPLICATION OF

PIPPIG ET AL.

APPLICATION NO:

FILED:

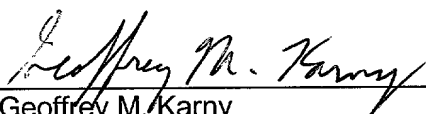
FOR: SELECTABLE CELL SURFACE MARKER GENES

Commissioner for Patents  
Washington, D.C. 20231SUBMISSION OF SEQUENCE LISTING  
INCLUDING STATEMENT OF VERIFICATION

Sir:

Applicants hereby provide a Computer Readable Form of the Sequence Listing as well as the Paper Copy thereof. The undersigned states that the Paper Copy and the Computer Readable Form, submitted in accordance with 37 CFR §1.821(c) and (e), respectively, are the same.

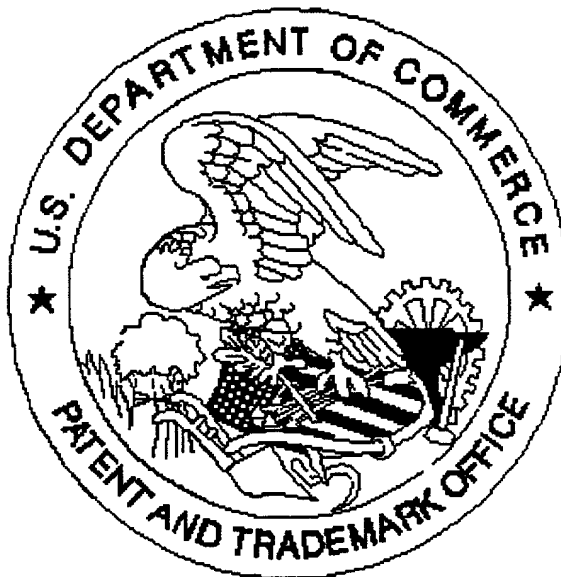
Respectfully submitted,

Novartis Corporation  
Patent and Trademark Dept.  
564 Morris Avenue  
Summit, NJ 07901-1027  
\_\_\_\_\_  
Geoffrey M. Karny  
Attorney for Applicants  
Reg. No. 31,382  
(301) 258-4640

Date: November 17, 2000

jc860 U.S. PTO  
09/715249  
11/17/00

United States Patent & Trademark Office  
Office of Initial Patent Examination -- Scanning Division



Application deficiencies found during scanning:

☐ Page(s) \_\_\_\_\_ of \_\_\_\_\_ were not present  
for scanning. (Document title)

☐ Page(s) \_\_\_\_\_ of \_\_\_\_\_ were not present  
for scanning. (Document title)

Missing page 41 of Specification

☐ **Scanned copy is best available.**